JC07 Rec'd PCT/PTO 0 4 DEC 2001

FORM PTO-1390

U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER PF-0709 USN

U.S APPLICATION NO. OKINGING 32918

INTERNATIONAL APPLICATION NO PCT/US00/16668

INTERNATIONAL FILING DATE 16 June 2000

PRIORITY DATE CLAIMED
17 June 1999

TITLE OF INVENTION

HUMAN TRANSPORT PROTEINS

APPLICANT(S) FOR DO/EO/US

INCYTE GENOMICS, INC.; LAL, Freeti; YANG, Junming; YUE, Henry, HILLMAN, Jennifer L.; TANG, Y. Tom; BANDMAN, Olga; BURIORD, Neil; BAUGHN, Mariah R.; AZIMZAI, Yalda; LU, Dyung Aina M.; AU-YOUNG, Janice; PATTERSON, Chandra

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

- 1. This is the FIRST submission of items concerning a filing under 35 U.S.C. 371.
- 2.

 This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
- 3.

 This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
- 4.
 ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
- 5.

 A copy of the International Application as filed (35 U.S.C. 371(c)(2))
 - a. \square is attached hereto (required only if not communicated by the International Bureau)
 - b. \square has been communicated by the International Bureau.
 - c. Sis not required, as the application was filed in the United States Receiving Office (RO/US).
- 6. □ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
- - a. \square are attached hereto (required only if not communicated by the International Bureau).
 - b. \square have been communicated by the International Bureau.
 - c. \square have not been made; however, the time limit for making such amendments has NOT expired.
- 8. □ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
- 9.

 ✓ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
- 10.□ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

- 11. □ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
- 12. An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
- 13. A FIRST preliminary amendment, as follows: Cancel in this application original claims 12, 14, 18, 20, 21, 23, 24, 25, 28 230 before calculating the filing fee, without prejudice or disclaimer. Applicants submit that these claims were included in the application as filed in the interest of providing notice to the public of certain specific subject matter intended to be claimed, and are being canceled at this time in the interest of reducing filing costs. Applicants expressly state that these claims are not being canceled for reasons related to patentability, and are in fact fully supported by the specification as filed. Applicants expressly reserve the right to reinstate these claims or to add other claims during prosecution of this application or a continuation or divisional application. Applicants expressly do not disclaim the subject matter of any invention disclosed herein which is not set forth in the instantly filed claims.
 - ☐ A SECOND or SUBSEQUENT preliminary amendment.
- 14. □ A substitute specification.
- 15. \square A change of power of attorney and/or address letter.
- 16.

 Other items or information:
- 1) Transmittal Letter (2 pp, in duplicate)
- 2) Return Postcard
- 3) Express Mail Label No.: EL 697 344 232 US
- 4) Article 34 Amendment

JC07 Rec'd PCT/PTO 0 4 DEC 2001

		W-1,	200			4 DLG 2001	
U.S. APPLICATION NO (III GOVED S Q 373 FR 2.58		INTERNATIONAL APP PCT/US00/16668	INTERNATIONAL APPLICATION NO PCT/US00/16668		ATTORNEY'S DOCKET NUMBER PF-0709 USN		
Neither internation nor internationalse. and International Sc □International preli USPTO but Interna International prelim but international pre ■International pre but all claims did i □International prelim	FEE (37 CFR 1.492(a)(1) alpreliminary examination arch fee (37 CFR 1.445(a)) earch Report not prepared minary examination fee (3 tional Search Report prepared minary examination fee (37 cFR 1 445(a)(a)(a)(a)(a)(b) provisions of Iminary examination fee materials and satisfy provisions of Iminary examination fee materials and satisfy provisions of Iminary examination fee materials are satisfied as a second satisfy provisions of Iminary examination fee materials are satisfied as a second satisfy provisions of Iminary examination fee materials are satisfied as a second satisfy provisions of Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision so Iminary examination fee materials are satisfied as a second satisfy provision satisfy provision satisfy provision satisfy provision satisfy provision satisf	fee (37 CFR 1.482) (2)) paid to USPTO bythe EPO or JPO\$100	50 00 PTO 10 00 1.1482) \$710.00				
	ENTER APPROPRIATE BASIC FEE AMOUNT =						
Surcharge of \$130.00 for furnishing the oath or decaration later than \square 20 \square 30 months from the earliest claimed priority date (37 CFR 1.492(e)).					\$		
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE				
Total Claims	20 =	0	X \$ 18.00		\$		
Independent Chims	3 =	1	X \$ 80.00		\$ 80.00		
MULTIPLE DEPEND	ENT CLAIM(S) (if applica	ible)	+ \$270.00		\$		
TOTAL OF ABOVE CALCULATIONS =					\$790.00		
□ Applicant claims small entity status. See 37 CFR 1 27. The fees indicated above are reduced by 1/2.					\$		
SUBTOTAL =					\$790.00		
Processing fee of \$130.00 for furnishing the English translation later than 20 30 months from the earliest clailmed priority date (37 CFR 1492(f)).					\$		
TOTAL NATIONAL FEE =					\$790.00		
Fee for recording the encbsed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property +					\$		
TOTAL FEES ENCLOSED =					\$790.00		
					Amount to be Refunded	\$	
					Charged	\$	
 c.	Deposit Account No. 09-0 is herebyauthorized to ch	to cover the above fees is 108 in the amount of \$790 arge anyadditional fees who 8. A duplicate copy of this	1.00 to cover the a	d or credit and	,		
NOTE: Where an appr be filed and granted to	opriate time limit under restore the application t	37 CFR 1.494 or 1.495 has o pending status.	s not been met, a	petition to rev	vive (37 CIR 1.1	137(a) or (b)) must	
SEND ALL CORRESPONDENCE TO: NCYTE GENOMICS, INC. 1160 Porter Drive Palo Alto, CA 94304							
		NAME: Diana Hamlet-C	Cox				
REGISTRATION NUMBER: 33,302							
		DATE: 4 De	cember 2001			İ	

10/009328 Docket No.: PF-0709 USN

"Express Mail" mailing label number <u>EL 697 344 232 US.</u> I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Commissioner for Patents, Box Patent Application, P.O. Box 2327, Arlington, VA 22202 on ——Decemb

By: Menu Roma Printed: NA

December 2001

Printed: 10 AND

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Preeti Lal, Junming Yang, Henry Yue, Jennifer L. Hillman, Y. Tom Tang, Olga Bandman, Neil Burford, Mariah R. Baughn, Yalda Azimzai, Dyung Aina M. Lu, Janice Au-Young & Chandra Patterson

Title:

HUMAN TRANSPORT PROTEINS

PCT Serial No.: PCT/US00/16668

International Filing Date: 16 June 2000

Examiner:

To Be Assigned

Group Art Unit:

To Be Assigned

Assistant Commissioner for Patents

Box Patent Application Washington, D.C. 20231

SUBMISSION UNDER 37 CFR § 1.821-1.825 SEQUENCE LISTING

Sir:

In accordance with the requirements of 37 CFR § 1.821-1.825, Applicants hereby submit one (1) diskette(s) containing the computer-readable information for the Sequence Listing of the above-identified application. The content of the Sequence Listing paper copy is identical to the computer-readable copy filed with the US Receiving Office. The USPTO is authorized to add whatever is necessary to update the CRF with the current application information.

Respectfully submitted,

Date: Poecember 2001

INCYTE GENOMICS, INC.

Diana Hamlet-Cox Reg. No. 33,302

Direct Dial Telephone: (650) 845-4639

3160 Porter Drive Palo Alto, California 94304 Phone: (650) 855-0555 Fax: (650) 845-4166

HUMAN TRANSPORT PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human transport proteins and to the use of these sequences in the diagnosis, treatment, and prevention of transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer.

5

10

20

30

35

BACKGROUND OF THE INVENTION

Eukaryotic cells are surrounded and subdivided into functionally distinct organelles by hydrophobic lipid bilayer membranes. These membranes act as a barrier to most molecules, and maintain the essential differences between the cytosol, the extracellular environment, and the contents of each intracellular organelle. Transport of essential nutrients, certain metal ions, metabolic waste products, cell signaling molecules, macromolecules, and proteins across lipid membranes and between organelles must be mediated by a variety of transport molecules. Transport between the cytoplasm and the extracellular environment, and between the cytoplasm and lumenal spaces of cellular organelles requires specific transport proteins. Each transport protein carries a particular class of molecule, such as ions, sugars, or amino acids, and often is specific to a certain molecular species of the class.

Cells and organelles require transport molecules to import and export essential nutrients and metal ions including K⁺, NH₄⁺, P₁, SO₄²⁻, sugars, and vitamins, as well as various metabolic waste products. Transport proteins also play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). Transport can occur by a passive concentration-dependent mechanism, or can be linked to an energy source such as ATP hydrolysis or an ion gradient. Proteins that function in transport include carrier proteins, which bind to a specific solute and undergo a conformational change that transfers the bound solute across the membrane, and channel proteins, which form hydrophilic pores that allow specific solutes to diffuse through the membrane down an electrochemical solute gradient.

Transport proteins are often multi-pass transmembrane proteins, which either actively transport molecules across the membrane or passively allow them to cross. Active transport involves directional pumping of a solute across the membrane, usually against an electrochemical gradient. Active transport is tightly coupled to a source of metabolic energy, such as ATP hydrolysis or an electrochemically favorable ion gradient. Passive transport involves the movement of a solute down

its electrochemical gradient. Transport proteins can be further classified as either carrier proteins or channel proteins. Carrier proteins, which can function in active or passive transport, bind to a specific solute to be transported and undergo a conformational change which transfers the bound solute across the membrane. Channel proteins, which only function in passive transport, form hydrophilic pores across the membrane. When the pores open, specific solutes, such as inorganic ions, pass through the membrane and down the electrochemical gradient of the solute. Examples include facilitative transporters, the secondary active symporters and antiporters driven by ion gradients, and active ATP binding cassette transporters involved in multiple-drug resistance and targeting of antigenic peptides to MHC Class I molecules. Transported substrates range from nutrients and ions to a broad variety of drugs, peptides and proteins.

Information on the action of ARL-6 (ADP-ribosylation like factor), an endoplasmic reticulum transmembrane protein, can be found in Greenfield, J.J. and S. High (1999; J. Cell Sci. 112:1477-1486). Information on reduced folate carrier transporter proteins can be found in Dixon, K.H. et al. (1994; J. Biol. Chem. 269:17-20) and Moscow, J.A. et al. (1995; Cancer Res. 55:5983-5987).

10

15

20

25

30

35

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelia contain a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous Na⁺/K⁺ ATPase. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N-and C-termini.

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions and charged metabolites between the cytosol and the mitochondrial matrix. Examples include the ADP, ATP carrier protein; the 2-oxoglutarate/malate carrier; the phosphate carrier protein; the brown fat uncoupling protein which transports protons from the cytosol into the matrix; the pyruvate carrier; the dicarboxylate carrier which transports malate, succinate, fumarate, and phosphate; the tricarboxylate carrier which transports citrate and malate; and the Grave's disease carrier protein, a protein recognized by IgG in patients with active Grave's disease, an autoimmune disorder resulting in hyperthyroidism (Stryer, L. (1995) <u>Biochemistry</u>, W.H. Freeman and Company, New York NY, p. 551; PROSITE PDOC00189 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) *275000 Graves Disease).

This class of transporters also includes the mitochondrial uncoupling proteins, which create

proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. The result is energy dissipation in the form of heat. Mitochondrial uncoupling proteins have been implicated as modulators of thermoregulation and metabolic rate, and have been proposed as potential targets for drugs against metabolic diseases such as obesity (Ricquier, D. et al. (1999) J. Int. Med. 245:637-642).

A number of metal ions such as iron, zinc, copper, cobalt, manganese, molybdenum, selenium, nickel, and chromium are important as cofactors for a number of enzymes. For example, zinc is required for the function of enzymes such as the extracellular matrix metalloproteinases, and zinc ions stabilize several motifs commonly found in transcription factors, including zinc fingers, zinc clusters, and LIM domains. Zinc and other metal ions must be provided in the diet, and are absorbed by transporters in the gastrointestinal tract. Plasma proteins transport the metal ions to the liver and other target organs, where specific transporters move the ions into cells and cellular organelles as needed. Imbalances in metal ion metabolism have been associated with a number of disease states (Danks, D.M. (1986) J. Med. Genet. 23:99-106).

10

15

20

30

35

The largest and most diverse family of transport proteins known are the ATP-binding cassette (ABC) transporters. As a family, ABC transporters can transport substances that differ markedly in chemical structure and size, ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC proteins consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) Methods Enzymol. 292:131-163).

The nuclear pore complex (NPC) is a large multiprotein complex spanning the nuclear envelope which mediates the transport of proteins and RNA molecules between the nucleus and the cytoplasm, thus contributing to the regulation of gene expression. The NPC allows passive diffusion of ions, small molecules, and macromolecules under about 60kD, while larger macromolecules are transported by facilitated, energy-dependent pathways. Nuclear localization signals (NLS), consisting

of short stretches of amino acids enriched in basic residues, are found on proteins that are targeted to the nucleus, such as the glucocorticoid receptor. The NLS is recognized by the NLS receptor, importin, which then interacts with the monomeric GTP-binding protein Ran. This NLS protein/receptor/Ran complex navigates the nuclear pore with the help of the homodimeric protein nuclear transport factor 2 (NTF2) (Nakielny, S. and G. Dreyfuss (1997) Curr. Opin. Cell Biol. 9:420-429; Gorlich, D. (1997) Curr. Opin. Cell Biol. 9:412-419). Four O-linked glycoproteins, p62, p58, p54, and p45, exist as a stable "p62 complex" that forms a ring localized on both nucleoplasmic and cytoplasmic surfaces of the NPC. The p62, p58, and p54 proteins all interact directly with the cytosolic transport factors p97 and NTF2, suggesting that the p62 complex is an important ligand binding site near the central gated channel of the NPC (Hu, T. et al. (1996) J. Cell Biol. 134:589-601).

Transport can also occur through intercellular bridges which connect the cytoplasms of sister cells, for example in the male and female germline of species ranging from fruit flies to humans. These bridges allow passage of cytoplasmic materials between cells during development.

Intercellular bridges have also been found to connect somatic cells. The nurse cells and oocyte of a Drosophila egg chamber, which are derived from a single precursor cell through four rounds of mitosis, are connected to each other through intercellular bridges called ring canals. The cells do not completely separate after mitosis; the mitotic cleavage furrows are transformed into ring canals by the addition of an actin cytoskeleton lining the tunnels between the cells. The Drosophila kelch protein functions in organizing actin in the ring canal. Mutations in kelch cause female sterility in Drosophila. Kelch contains four protein domains: the NTR domain at the N-terminus, the BTB or POZ domain, the IVR or intervening region; and the kelch repeat domain, which contains six 50-amino acid kelch repeats. The BTB or POZ domain, a 120-amino acid motif that is also found in several zinc-finger containing transcription factors, may be important in dimerization of kelch. Kelch repeats are found in other proteins as well and may be important for actin binding (Robinson, D.N. and L. Cooley (1997) J. Cell Biol. 138:799-810; Cooley, L. (1998) Cell 93:913-915).

Ion Channels

10

15

20

25

30

35

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form an ion-selective pore within the membrane. Ion channels share common structural and mechanistic themes. The channel consists of four or five subunits or protein monomers that are arranged like a barrel in the plasma membrane. Each subunit typically consists of six potential transmembrane segments (S1, S2, S3, S4, S5, and S6). The center of the barrel forms a pore lined by α -helices or β -strands. The side chains of the amino acid residues comprising the α -helices or β -strands establish the charge (cation or anion) selectivity of the channel. The degree of selectivity, or what specific ions are allowed to pass through the channel, depends on the diameter of the narrowest part of the pore. There

are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

Transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including Na⁺-K⁺ ATPase, Ca²⁺ ATPase, H⁺ ATPase, and Cu⁺⁺ ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na⁺ and Ca²⁺ are low and cytosolic concentration of K⁺ is high. The vacuolar (V) class of ion transporters include H⁺ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H⁺ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate (P_i).

10

25

35

Cu⁺⁺ ATPases export copper from cells (PROSITE PDOC00139 E1-E2 ATPases phosphorylation site). Mutations in one Cu⁺⁺ ATPase cause Wilson disease, in which toxic amounts of copper accumulate in a number of organs, particularly the liver and brain (Tanzi, R.E. et al. (1993) Nat. Genet. 5:344-350). Mutations in another Cu⁺⁺ ATPase cause Menkes disease and occipital horn syndrome. Menkes disease mutations block export of copper from the gastrointestinal tract, leading to skeletal abnormalities, severe mental retardation, neurologic degeneration, and mortality in early childhood (Harrison, M.D. and C.T. Dameron (1999) J. Biochem. Mol. Toxicol. 13:93-106). Occipital horn syndrome mutations cause connective tissue defects (Harrison, supra; Levinson, B. et al. (1996) Hum. Mol. Genet. 5:1737-1742).

The coupling factor (F) class of ion transporters consists of H⁺ pumps in mitochondria, chloroplasts, and bacteria. For example, the F_0F_1 ATPase utilizes a proton gradient across the inner mitochondrial membrane to generate ATP from ADP and inorganic phosphate (P₁). The F_0F_1 ATPase is composed of the F_0 complex, which is the transmembrane channel through which protons flow, and the F_1 complex, where ATP synthesis activity resides. F_0 has three subunits, A (also known as protein 6), B, and C (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 752-756; PROSITE PDOC00420 ATP synthase a subunit signature).

Voltage-gated Ca²⁺ channels are involved in presynaptic neurotransmitter release, and heart and skeletal muscle contraction. The voltage-gated Ca²⁺ channels from skeletal muscle (L-type) and brain (N-type) have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The α_1 subunit forms the

membrane pore and voltage sensor, while the $\alpha_2\delta$ and β subunits modulate the voltage-dependence, gating properties, and the current amplitude of the channel. These subunits are encoded by at least six α_1 , one $\alpha_2\delta$, and four β genes. A fourth subunit, γ , has been identified in skeletal muscle (Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; and Jay, S.D. et al. (1990) Science 248:490-492). The human $\beta 4$ subunit is homologous to the mouse epilepsy gene <u>lethargic</u>, and is a candidate for involvement in neurological disorders including ataxia and absence epilepsy (Escayg, A. et al. (1998) Genomics 50:14-22).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of Na⁺ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as γ -aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential.

10

15

20

25

30

Ion channels are expressed in a number of tissues where they are implicated in a variety of processes. CNG channels, while abundantly expressed in photoreceptor and olfactory sensory cells, are also found in kidney, lung, pineal, retinal ganglion cells, testis, aorta, and brain. Calcium-activated K+ channels may be responsible for the vasodilatory effects of bradykinin in the kidney and for shunting excess K+ from brain capillary endothelial cells into the blood. They are also implicated in repolarizing granulocytes after agonist-stimulated depolarization (Ishi, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656). Another transmembrane protein, the leukotrine B4 receptor (BLT) appears to be involved in inflammation responses and host cell defense against infection. BLT also functions as an HIV coreceptor (Izumi, T. et al. (1997) Nature 387:620-624; Martin, V. et al. (1999) J. Biol. Chem. 274:8597-8603).

Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98).

K⁺ channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as Ca⁺⁺ and cAMP. In non-excitable tissue, K⁺ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K⁺ channels are responsible for setting resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell

contains a Na⁺-K⁺ pump and ion channels that provide the redistribution of Na⁺, K⁺, and Cl⁻. The pump actively transports Na⁺ out of the cell and K⁺ into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K⁺ and Cl⁻ to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl⁻ flows out of the cell. The flow of K⁺ is balanced by an electromotive force pulling K⁺ into the cell, and a K⁺ concentration gradient pushing K⁺ out of the cell. Thus, the resting membrane potential is primarily regulated by K⁺ flow (Salkoff, L. and T. Jegla (1995) Neuron 15:489-492). Information on NY-REN-45, a K+ channel integral membrane protein, can be found in Scanlan, M.J. et al. (1998; Int. J. Cancer 76:652-658). The emopamil-binding protein (EBP) shares structural features with both pro- and eukaryotic drug transport proteins (Hanner, M. et al. (1995) J. Biol. Chem. 270:7551-7557). The Na+ channel, transmembrane protein myelin protein zero (MPZ) may be responsible for some sporadic cases of Dejerine-Scottas disease (hereditary motor and sensory neuropathy type III) (Hayasaka, K. et al. (1993) Nat. Genet. 5:266-268).

10

15

20

25

30

35

K⁺ pore-forming subunits generally have six transmembrane-spanning domains with a short region between the fifth and sixth transmembrane regions that senses membrane potential; and the amino and carboxy termini are located intracellularly. In mammalian heart, the duration of ventricular action potential is controlled by a K⁺ current. Thus, the K⁺ channel is central to the control of heart rate and rhythm. K⁺ channel dysfunctions are associated with a number of renal diseases including hypertension, hypokalemia, and the associated Bartter's syndrome and Getelman's syndrome, as well as neurological disorders including epilepsy. K⁺ channels have been implicated in Alzheimer's disease by observations that a significant component of senile plaques, beta amyloid or A beta, also blocks voltage-gated potassium channels in hippocampal neurons (Antes, L.M. et al. (1998) Seminar Nephrol. 18:31-45; Stoffel, M. and L.Y. Jan (1998) Nat. Genet. 18:6-8; Madeja, M. et al. (1997) Eur. J. Neurosci. 9:390-395; Good, T.A. et al. (1996) Biophys. J. 70:296-304).

Gated ion channels control ion flow by regulating the opening and closing of pores. These channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open pores in response to mechanical stress, voltage-gated channels open pores in response to changes in membrane potential, and ligand-gated channels open pores in the presence of a specific ion, nucleotide, or neurotransmitter.

Voltage-gated Na⁺ channels are responsible for electrical excitability of neurons, skeletal muscle, heart, and neuroendocrine tissues. For example, the sequential opening and closing of voltage-gated Na⁺ channels results in the propagation of action potentials down neuronal axons. Na⁺ channels isolated from rat brain tissue are heterotrimeric complexes composed of a 260 kDa pore forming α subunit that associates with two smaller auxiliary subunits, β 1 and β 2. The β 2 subunit is an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with α and β 1 subunits correlates with increased function of the channel, a change in the channel's gating properties, as well as an increase in whole cell capacitance (Isom, L.L. et al. (1995) Cell 83:433-442).

Integral Membrane Proteins

20

25

30

35

The majority of known integral membrane proteins are transmembrane proteins (TM) which are characterized by an extracellular, a transmembrane, and an intracellular domain. TM domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an α-helical conformation. TM proteins are classified as bitopic (Types I and II) and polytopic (Types III and IV) (Singer, S.J. (1990) Annu. Rev. Cell Biol. 6:247-96). Bitopic proteins span the membrane once while polytopic proteins contain multiple membrane-spanning segments. TM proteins that act as cell-surface receptor proteins involved in signal transduction include growth and differentiation factor receptors, and receptor-interacting proteins such as Drosophila pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane proteins. TM proteins also act as transporters of ions or metabolites, such as gap junction channels (connexins) and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TM proteins act as vesicle organelle-forming molecules, such as calveolins, or as cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins. Information on connexin can be found in Kanter, H.L. et al. (1994; J. Mol. Cell. Cardiol. 26:861-868).

Many membrane proteins (MPs) contain amino acid sequence motifs that target these proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in cancer treatments which target tumor vasculature (Arap, W. et al. (1998) Science, 279:377-380.) Furthermore, MPs may also contain amino acid sequence motifs, such as the carbohydrate recognition domain (CRD), also known as the C-type lectin domain, that mediate interactions with extracellular or intracellular molecules.

G-protein coupled receptors (GPCR) comprise a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines, lipid mediators of inflammation, peptide hormones, and sensory signal mediators. The structure of these highly-conserved receptors consists of seven hydrophobic transmembrane regions, an extracellular N-terminus, and a cytoplasmic C-terminus. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. Cysteine disulfide bridges connect the second and third extracellular loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with G proteins. A GPCR consensus pattern is characteristic of most proteins belonging to this superfamily (ExPASy PROSITE document PS00237; and Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp 2-6). Mutations and changes in transcriptional activation of GPCR-encoding genes have been

associated with neurological disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, drug addiction, and feeding disorders.

Cytochromes are electron-transferring proteins that contain a heme prosthetic group, a porphyrin ring containing a tightly bound iron atom. Cytochromes act as oxidoreductases in such diverse cellular processes as respiration, photosynthesis, fatty acid metabolism, and neurotransmitter biosynthesis. The heme iron atom serves as the actual electron carrier by changing from the ferric to the ferrous oxidation state when accepting an electron. Cytochromes accept electrons from one substrate such as NADH or ascorbate and donate them to other electron carriers such as other cytochromes, ubiquinone, or semidehydroascorbic acid (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 759-770, 786-797; Sperling, P. et al. (1995) Eur. J. Biochem. 232:798-805; and Online Mendelian Inheritance in Man (OMIM) *600019 Cytochrome b561, CYB561).

Cytochrome b5 is an electron donor in membrane-linked redox enzyme systems involved in lipid and drug metabolism. Cytochrome b5 has been found in Golgi, plasma, outer mitochondrial, endoplasmic reticulum (ER), and microbody membranes. Conserved amino acids in cytochrome b5 15 include eight invariant amino acids at W34, H51, P52, G53, G54, G63, F70, and H74, and fifteen conserved amino acids at L24, I35, S36, V41, Y42, N43, T45, W47, A48, L58, D65, T67, L85, T87, and G88 (numbering based on the sunflower cytochrome b5/delta-6 desaturase fusion protein; GI 1040729, Sperling, supra). The invariant residues H51PGG are involved in heme-binding. Cytochrome b5-like domains have also been found linked to other enzymes. For example, cytochrome b5-like domains are part of delta-9 fatty acid desaturases in yeast and Histoplasma capsulatum, nitrate reductase, sulfite reductase, flavocytochrome b2, Arabidopsis thaliana acyl lipid desaturase, and Borago officinalis (borage) and Helianthus annuus (sunflower) delta-6 desaturases (Sperling, supra; Sayanova, O. et al (1997) Proc. Natl. Acad. Sci. USA 94:4211-4216; and Mitchell, A.G. and C.E. Martin (1997) J. Biol. Chem. 272:28281-28288).

Signal peptides are found on proteins that are targeted to the endoplasmic reticulum (ER). Signal peptides consist of stretches of amino acids enriched in hydrophobic residues. Signal peptides are usually found at the extreme N-terminus of the protein and are recognized by a cytosolic signalrecognition peptide (SRP). The SRP binds to the signal peptide and to an SRP receptor, an integral membrane protein in the ER. Once bound to the SRP receptor, the newly formed protein containing the signal peptide is translocated across the ER membrane. Proteins containing signal peptides may end up inserted into the lipid bilayer, or they may end up in the lumen of an organelle or secreted from the cell.

35 **Disease Correlation**

20

25

The etiology of numerous human diseases and disorders can be attributed to defects in the

transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g. cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small molecules across membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) Exp. Nephrol. 4:253-262; Talente, G.M. et al. (1994) Ann. Intern. Med. 120:218-226; and Chillon, M. et al. (1995) New Engl. J. Med. 332:1475-1480).

Cystinuria is an inherited disease that results from the inability to transport cystine, the disulfide-linked dimer of cysteine, from the urine into the blood. Accumulation of cystine in the urine leads to the formation of cystine stones in the kidneys.

10

15

20

25

30

35

Transthyretin (TTR), present in human plasma, binds to and transports the thyroid hormone thyroxine. Mutations in TTR result in the conversion of TTR to amyloid, an insoluble fibrillar structure. The resulting amyloid plaques have been shown to be the causative agent in the development of familial amyloid polyneuropathy and senile systemic amyloidosis (Miroy, G.J. et al. (1996) Proc. Natl. Acad. Sci. USA 93:15051-15056).

Stomatin, a 31-kDa erythrocyte integral membrane protein has been linked to the hereditary anemia stomatocytosis. This anemia is characterized by red blood cells that lack stomatin and leak Na+ and K+. Thus, stomatin is presumed to play a role in the regulation of ion transport. Red blood cell ion transport defects are also linked to other disorders such as hypertension (Stewart, G.W. (1997) Int. J. Biochem. Cell Biol. 29:271-274).

The discovery of new human transport proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, human transport proteins, referred to collectively as "TPPT" and individually as "TPPT-1," "TPPT-2," "TPPT-3," "TPPT-4," "TPPT-5," "TPPT-6," "TPPT-7," "TPPT-8," "TPPT-9," "TPPT-10," "TPPT-11," "TPPT-12," "TPPT-13," "TPPT-14," "TPPT-15," "TPPT-16," "TPPT-17," "TPPT-18," "TPPT-19," "TPPT-20," "TPPT-21," "TPPT-22," "TPPT-23," "TPPT-24," "TPPT-25," "TPPT-26," "TPPT-27," "TPPT-28," "TPPT-29," "TPPT-30," "TPPT-31," "TPPT-32," "TPPT-33," "TPPT-34," "TPPT-35," "TPPT-36," "TPPT-37," "TPPT-38," "TPPT-39," "TPPT-40," "TPPT-41," "TPPT-42," and "TPPT-43." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-

43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-43.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-43. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:44-86.

10

15

20

25

30

35

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a

polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

10

15

20

25

30

35

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and a pharmaceutically acceptable excipient. In one embodiment, the pharmaceutical composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

10

15

20

25

30

35

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional TPPT, comprising administering to a patient in

need of such treatment the pharmaceutical composition.

15

20

25

30

35

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:44-86, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding TPPT.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of TPPT.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding TPPT were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

10

15

20

25

35

5

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

30 **DEFINITIONS**

"TPPT" refers to the amino acid sequences of substantially purified TPPT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of TPPT. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TPPT either by directly interacting with

TPPT or by acting on components of the biological pathway in which TPPT participates.

10

20

25

30

35

An "allelic variant" is an alternative form of the gene encoding TPPT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding TPPT include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TPPT or a polypeptide with at least one functional characteristic of TPPT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TPPT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding TPPT. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TPPT. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of TPPT is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of TPPT. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TPPT either by directly interacting with TPPT or by acting on components of the biological pathway in which TPPT

participates.

10

15

20

25

30

35

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind TPPT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic TPPT, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or

amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding TPPT or fragments of TPPT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution		
20	Ala	Gly, Ser		
	Arg	His, Lys		
	Asn	Asp, Gln, His		
	Asp	Asn, Glu		
	Cys	Ala, Ser		
25	Gln	Asn, Glu, His		
	Glu	Asp, Gln, His		
	Gly	Ala		
	His	Asn, Arg, Gln, Glu		
	Ile	Leu, Val		
30	Leu	Ile, Val		
	Lys	Arg, Gln, Glu		
	Met	Leu, Ile		
	Phe	His, Met, Leu, Trp, Tyr		
	Ser	Cys, Thr		
35	Thr	Ser, Val		
	Trp	Phe, Tyr		
	Tyr	His, Phe, Trp		
	Val	Ile, Leu, Thr		

15

40

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

10

15

20

25

30

35

A "fragment" is a unique portion of TPPT or the polynucleotide encoding TPPT which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:44-86 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:44-86, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:44-86 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:44-86 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:44-86 and the region of SEQ ID NO:44-86 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-43 is encoded by a fragment of SEQ ID NO:44-86. A fragment of SEQ ID NO:1-43 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-43. For example, a fragment of SEQ ID NO:1-43 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-43. The precise length of a fragment of SEQ ID NO:1-43 and the region of SEQ ID NO:1-43 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended

purpose for the fragment.

5

10

15

20

25

30

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at

http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version

2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example: *Matrix: BLOSUM62*

35 Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

5 Filter: on

15

20

25

30

35

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10
Word Size: 3
Filter: on

15

20

25

30

35

Percent identity may be measured over the length of an entire defined polypeptide sequence,

for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for
example, over the length of a fragment taken from a larger, defined polypeptide sequence, for
instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least
150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment
length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be
used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0 t or R_0 t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

10

15

20

25

30

35

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TPPT which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of TPPT which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of TPPT. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of TPPT.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide,

polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

5

10

15

20

25

30

35

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an TPPT may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TPPT.

"Probe" refers to nucleic acid sequences encoding TPPT, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs

can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

15

20

25

30

35

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, <u>supra</u>. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is

expressed, inducing a protective immunological response in the mammal.

5

10

15

20

25

30

35

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding TPPT, or fragments thereof, or TPPT itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

10

20

25

35

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide

polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

10 THE INVENTION

15

20

25

35

The invention is based on the discovery of new human transport proteins (TPPT), the polynucleotides encoding TPPT, and the use of these compositions for the diagnosis, treatment, or prevention of transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding TPPT. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each TPPT were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide sequence of each TPPT and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding TPPT. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:44-86

and to distinguish between SEQ ID NO:44-86 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express TPPT as a fraction of total tissues expressing TPPT. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing TPPT as a fraction of total tissues expressing TPPT. Column 5 lists the vectors used to subclone each cDNA library.

Of particular interest is the expression of SEQ ID NO:50 exclusively in cardiovascular tissue, the expression of SEQ ID NO:56 in nervous and gastrointestinal tissues, the expression of SEQ ID NO:57 in gastrointestinal tissues, and the expression of SEQ ID NO:66 in nervous system tissues. Of particular note is the tissue-specific expression of SEQ ID NO:75. Over 71% of the cDNA libraries expressing SEQ ID NO:75 are derived from lung tissue.

10

15

25

30

35

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding TPPT were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:44 maps to chromosome 7 within the interval from 38.80 to 42.10 centiMorgans. SEQ ID NO:48 maps to chromosome X within the interval from 107.90 to 122.80 centiMorgans. SEQ ID NO:60 maps to chromosome 2 within the interval from 157.0 to 167.0 centiMorgans. SEQ ID NO:65 maps to chromosome 2 within the interval from 17.4 to 40.7 centiMorgans and to chromosome 5 within the interval from 61.1 to 69.6 centiMorgans. The interval on chromosome 5 from 61.1 to 69.6 centiMorgans also contains genes associated with Cockayne syndrome. SEQ ID NO:69 maps to chromosome 3 within the interval from 157.40 to 162.00 centiMorgans. SEQ ID NO:70 maps to chromosome 3 within the interval from 176.40 to 179.80 centiMorgans. SEQ ID NO:71 maps to chromosome 18 within the interval from the p-terminus to 52.30 centiMorgans. SEQ ID NO:73 maps to chromosome 17 within the interval from 75.70 to 84.20 centiMorgans, and to chromosome 2 within the interval from 204.70 to 209.30 centiMorgans. SEQ ID NO:76 maps to chromosome 20 within the interval from 79.00 to 94.40 centiMorgans. SEQ ID NO:80 maps to chromosome 18 within the interval from 1.60 to 6.20 centiMorgans, and to chromosome 11 within the interval from 117.90 to 126.00 centiMorgans. SEQ ID NO:83 maps to chromosome 17 within the interval from 67.60 to 69.30 centiMorgans, and from 83.8 centiMorgans to the q-terminus, and to chromosome 7 within the interval from 105.20 to 114.50 centiMorgans.

The invention also encompasses TPPT variants. A preferred TPPT variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the TPPT amino acid sequence, and which contains at least one functional or structural characteristic of TPPT.

The invention also encompasses polynucleotides which encode TPPT. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected

5

10

15

20

25

30

35

from the group consisting of SEQ ID NO:44-86, which encodes TPPT. The polynucleotide sequences of SEQ ID NO:44-86, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding TPPT. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TPPT. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:44-86 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:44-86. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TPPT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding TPPT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring TPPT, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode TPPT and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring TPPT under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TPPT or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TPPT and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode TPPT and TPPT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding TPPT or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:44-86 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, 10 Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 20 856-853.)

The nucleic acid sequences encoding TPPT may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo

30

Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

10

15

20

25

30

35

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode TPPT may be cloned in recombinant DNA molecules that direct expression of TPPT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express TPPT.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TPPT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or

improve the biological properties of TPPT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding TPPT may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, TPPT itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of TPPT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

25

30

35

In order to express a biologically active TPPT, the nucleotide sequences encoding TPPT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding TPPT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TPPT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TPPT and its initiation codon and upstream regulatory sequences are inserted into

the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding TPPT and appropriate transcriptional and translational control elements. These methods include <u>in vitro</u> recombinant DNA techniques, synthetic techniques, and <u>in vivo</u> genetic recombination. (See, e.g., Sambrook, J. et al. (1989) <u>Molecular Cloning, A Laboratory Manual</u>, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) <u>Current Protocols in Molecular Biology</u>, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

10

35

A variety of expression vector/host systems may be utilized to contain and express sequences encoding TPPT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) 25 Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al., (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding TPPT. For example, routine cloning,

subcloning, and propagation of polynucleotide sequences encoding TPPT can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TPPT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TPPT are needed, e.g. for the production of antibodies, vectors which direct high level expression of TPPT may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of TPPT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast <u>Saccharomyces cerevisiae</u> or <u>Pichia pastoris</u>. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, <u>supra</u>; Bitter, <u>supra</u>; and Scorer, <u>supra</u>.)

10

15

20

25

30

35

Plant systems may also be used for expression of TPPT. Transcription of sequences encoding TPPT may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TPPT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses TPPT in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet.

15:345-355.)

10

25

For long term production of recombinant proteins in mammalian systems, stable expression of TPPT in cell lines is preferred. For example, sequences encoding TPPT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding TPPT is inserted within a marker gene sequence, transformed cells containing sequences encoding TPPT can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TPPT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding TPPT and that express TPPT may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or

chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of TPPT using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TPPT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

10

15

20

25

30

35

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding TPPT include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding TPPT, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding TPPT may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode TPPT may be designed to contain signal sequences which direct secretion of TPPT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for

post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the

American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TPPT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric TPPT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TPPT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, c-myc, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metalchelate resins, respectively. FLAG, c-myc, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TPPT encoding sequence and the heterologous protein sequence, so that TPPT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

15

20

25

30

35

In a further embodiment of the invention, synthesis of radiolabeled TPPT may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

TPPT of the present invention or fragments thereof may be used to screen for compounds that specifically bind to TPPT. At least one and up to a plurality of test compounds may be screened for specific binding to TPPT. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of TPPT, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) <u>Current Protocols in Immunology</u> 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which TPPT binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express TPPT, either as a secreted

protein or on the cell membrane. Preferred cells include cells from mammals, yeast, <u>Drosophila</u>, or <u>E. coli</u>. Cells expressing TPPT or cell membrane fractions which contain TPPT are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TPPT or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with TPPT, either in solution or affixed to a solid support, and detecting the binding of TPPT to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

10

15

20

25

30

35

TPPT of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of TPPT. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for TPPT activity, wherein TPPT is combined with at least one test compound, and the activity of TPPT in the presence of a test compound is compared with the activity of TPPT in the absence of the test compound. A change in the activity of TPPT in the presence of the test compound is indicative of a compound that modulates the activity of TPPT. Alternatively, a test compound is combined with an in vitro or cell-free system comprising TPPT under conditions suitable for TPPT activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of TPPT may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding TPPT or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred

to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding TPPT may also be manipulated <u>in vitro</u> in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding TPPT can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding TPPT is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TPPT, e.g., by secreting TPPT in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74). THERAPEUTICS

10

15

20

25

30

35

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TPPT and human transport proteins. In addition, the expression of TPPT is closely associated with neurological, cardiovascular, reproductive, gastrointestinal, and hematopoietic/immune tissues, and inflammation, cell proliferation, and cancer. Therefore, TPPT appears to play a role in transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer. In the treatment of disorders associated with increased TPPT expression or activity, it is desirable to decrease the expression or activity of TPPT. In the treatment of disorders associated with decreased TPPT expression or activity, it is desirable to increase the expression or activity of TPPT.

Therefore, in one embodiment, TPPT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer; cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline

myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis; neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia; and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, pentosuria phenylketonuria, and pseudovitamin Ddeficiency rickets; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins,

10

15

20

30

35

5

10

15

20

25

30

35

thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis,

pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing TPPT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT including, but not limited to, those described above.

10

15

20

25

30

35

In a further embodiment, a pharmaceutical composition comprising a substantially purified TPPT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of TPPT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT including, but not limited to, those listed above.

In a further embodiment, an antagonist of TPPT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TPPT. Examples of such disorders include, but are not limited to, those transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer described above. In one aspect, an antibody which specifically binds TPPT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express TPPT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TPPT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TPPT including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made

by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

5

10

15

20

25

30

35

An antagonist of TPPT may be produced using methods which are generally known in the art. In particular, purified TPPT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TPPT. Antibodies to TPPT may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with TPPT or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TPPT have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of TPPT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to TPPT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TPPT-specific single

chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing <u>in vivo</u> production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for TPPT may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between TPPT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering TPPT epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TPPT. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of TPPT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions.

The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple TPPT epitopes, represents the average affinity, or avidity, of the antibodies for TPPT. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular TPPT epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the TPPT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar

procedures which ultimately require dissociation of TPPT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons,

35 New York NY).

10

15

20

25

30

The titer and avidity of polyclonal antibody preparations may be further evaluated to

determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TPPT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding TPPT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TPPT. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TPPT. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

15

20

25

30

35

In another embodiment of the invention, polynucleotides encoding TPPT may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and Somia, N. (1997) Nature 389:239-242)), (ii)

express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated

cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as <u>Candida albicans</u> and <u>Paracoccidioides brasiliensis</u>; and protozoan parasites such as <u>Plasmodium falciparum</u> and <u>Trypanosoma cruzi</u>). In the case where a genetic deficiency in TPPT expression or regulation causes disease, the expression of TPPT from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in TPPT are treated by constructing mammalian expression vectors encoding TPPT and introducing these vectors by mechanical means into TPPT-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

10

20

30

35

Expression vectors that may be effective for the expression of TPPT include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). TPPT may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TPPT from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TPPT expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TPPT under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus cis-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

10

15

20

25

30

35

In the alternațive, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding TPPT to cells which have one or more genetic abnormalities with respect to the expression of TPPT. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TPPT to target cells which have one or more genetic abnormalities with respect to the expression of TPPT. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TPPT to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has

been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

10

15

20

25

30

35

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TPPT to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotech. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for TPPT into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TPPTcoding RNAs and the synthesis of high levels of TPPT in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of TPPT into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have

been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

5

10

15

20

25

30

35

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding TPPT.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding TPPT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2'O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding TPPT. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular

chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased TPPT expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding TPPT may be therapeutically useful, and in the treatment of disorders associated with decreased TPPT expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TPPT may be therapeutically useful.

10

15

20

35

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding TPPT is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding TPPT are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding TPPT. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5.932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved

using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such pharmaceutical compositions may consist of TPPT, antibodies to TPPT, and mimetics, agonists, antagonists, or inhibitors of TPPT.

10

15

20

25

30

35

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Pharmaceutical compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of pharmaceutical compositions may be prepared for direct intracellular delivery of macromolecules comprising TPPT or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TPPT or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell

culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys. or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example TPPT or fragments thereof, antibodies of TPPT, and agonists, antagonists or inhibitors of TPPT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

5

10

15

20

25

30

35

In another embodiment, antibodies which specifically bind TPPT may be used for the diagnosis of disorders characterized by expression of TPPT, or in assays to monitor patients being treated with TPPT or agonists, antagonists, or inhibitors of TPPT. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TPPT include methods which utilize the antibody and a label to detect TPPT in human body fluids

or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring TPPT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TPPT expression. Normal or standard values for TPPT expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to TPPT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TPPT expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

10

15

20

25

30

35

In another embodiment of the invention, the polynucleotides encoding TPPT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of TPPT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TPPT, and to monitor regulation of TPPT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TPPT or closely related molecules may be used to identify nucleic acid sequences which encode TPPT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding TPPT, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TPPT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:44-86 or from genomic sequences including promoters, enhancers, and introns of the TPPT gene.

Means for producing specific hybridization probes for DNAs encoding TPPT include the cloning of polynucleotide sequences encoding TPPT or TPPT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding TPPT may be used for the diagnosis of disorders

5

10

15

20

25

30

35

associated with expression of TPPT. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer; cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis; neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia; and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, pentosuria phenylketonuria, and pseudovitamin Ddeficiency rickets; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial

nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,

10

20

25

30

35

allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathycandidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, 10 polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, trauma, and hematopoietic cancer including 15 lymphoma, leukemia, and myeloma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, 20 bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding TPPT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from 25 patients to detect altered TPPT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding TPPT may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TPPT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding TPPT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

30

35

In order to provide a basis for the diagnosis of a disorder associated with expression of TPPT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding TPPT, under conditions suitable for hybridization or amplification.

Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

10

15

20

25

35

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TPPT may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding TPPT, or a fragment of a polynucleotide complementary to the polynucleotide encoding TPPT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding TPPT may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding TPPT are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are

detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of TPPT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

10

15

25

30

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for TPPT, or TPPT or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-

2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in <u>DNA Microarrays: A Practical Approach</u>, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding TPPT may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

10

20

25

35

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding TPPT on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, TPPT, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug

screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between TPPT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TPPT, or fragments thereof, and washed. Bound TPPT is then detected by methods well known in the art. Purified TPPT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TPPT specifically compete with a test compound for binding TPPT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TPPT.

In additional embodiments, the nucleotide sequences which encode TPPT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/139,923, U.S. Ser. No. 60/148,177, U.S. Ser. No. 60/149,357, and U.S. Ser. No. 60/162,287, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

10

15

20

25

30

35

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated

using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

5

10

20

25

30

35

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ

Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

10

15

20

25

30

35

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such

as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:44-86. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

15

20

25

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, <u>supra</u>, ch. 7; Ausubel, 1995, <u>supra</u>, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

BLAST Score x Percent Identity

5 x minimum {length(Seq. 1), length(Seq. 2)}

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding TPPT occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous,

reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of TPPT Encoding Polynucleotides

10

15

20

25

30

35

The cDNA sequences which were used to assemble SEQ ID NO:44-49 and SEQ ID NO:54-86 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:44-49 and SEQ ID NO:54-86 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

The genetic map locations of SEQ ID NO:44, SEQ ID NO:48, SEQ ID NO:60, SEQ ID NO:65, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:76, SEQ ID NO:80, and SEQ ID NO:83 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:65, SEQ ID NO:73, SEQ ID NO:80, and SEQ ID NO:83, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:65, SEQ ID NO:73, SEQ ID NO:80, and SEQ ID NO:83 were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention section where applicable. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VI. Extension of TPPT Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:44-86 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this

fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

10

15

20

25

35

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1% agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent <u>E. coli</u> cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, the polynucleotide sequences of SEQ ID NO:44-86 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:44-86 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

10

15

20

25

30

35

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers.

Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

<u>Tissue or Cell Sample Preparation</u>

10

15

35

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and 20 poly(A)+ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)+ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)+ RNA with 25 GEMBRIGHT kits (Incyte). Specific control poly(A)* RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85 °C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. 30 (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification

uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

20 <u>Hybridization</u>

5

10

15

25

30

35

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried. Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

10

15

20

30

35

Sequences complementary to the TPPT-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TPPT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of TPPT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TPPT-encoding transcript.

X. Expression of TPPT

Expression and purification of TPPT is achieved using bacterial or virus-based expression systems. For expression of TPPT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the trp-lac (tac) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the lac operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TPPT upon induction with isopropyl beta-Dthiogalactopyranoside (IPTG). Expression of TPPT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TPPT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, TPPT is synthesized as a fusion protein with, e.g., glutathione Stransferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from TPPT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified TPPT obtained by these methods can be used directly in the assays shown in Examples XI and XV.

30 XI. Demonstration of TPPT Activity

10

15

20

25

35

TPPT transport activity is assayed by measuring uptake of labeled substrates into <u>Xenopus</u> laevis oocytes. Oocytes at stages V and VI are injected with TPPT mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 1mM Na₂HPO₄, 5 mM Hepes, 3.8 mM NaOH, 50μg/ml gentamycin, pH 7.8) to allow expression of TPPT. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids,

sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with ³H, fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in Na⁺-free medium, measuring the incorporated label, and comparing with controls. TPPT activity is proportional to the level of internalized labeled substrate.

XII. Functional Assays

10

15

30

35

TPPT function is assessed by expressing the sequences encoding TPPT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser opticsbased technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of TPPT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TPPT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TPPT and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of TPPT Specific Antibodies

5

10

20

25

30

35

TPPT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the TPPT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TPPT activity by, for example, binding the peptide or TPPT to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring TPPT Using Specific Antibodies

Naturally occurring or recombinant TPPT is substantially purified by immunoaffinity chromatography using antibodies specific for TPPT. An immunoaffinity column is constructed by covalently coupling anti-TPPT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TPPT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TPPT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TPPT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TPPT is collected.

XV. Identification of Molecules Which Interact with TPPT

TPPT, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TPPT, washed, and any wells with labeled TPPT complex are assayed. Data obtained using different concentrations of TPPT are used to calculate values for the number, affinity, and association of TPPT with the candidate molecules.

Alternatively, molecules interacting with TPPT are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially

available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

TPPT may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table

	SPLNFET01), 264114H1 (HNT2AGT01), TONSNOT01), 747955R6 (BRAITUT01), (PENCNOT03), 3344943H1 8), 4028320H1 (BRAINOT23), (MCLRUNT01)	(THYRNOT10), 3536452F6	(BRAINOT12), 2084989H1), 2084989T6 (PANCNOT04), (BONRTUT01), 3109754F6), 3700647H1 (SININOT05)	, 414	MCINOTO1), 009314H1 (HMCINOTO1), SYNOOATO1), 1000842R1), 1374329H1 (BSTMMONO2), (TESTNOT11), SAIA02182F1	4	T01), SCDA08642V1	T03), 1715155F6 26143H1 (OVARNOT02), TUT16)	ADRENOT09), 2958321X305D1	ADTUT08), 3735780F6 3735780T6 (SMCCNOS01),	B01), 159164R1 (ADENINBO1)	1272850T1 F6 (KIDNNO'), 47993181
Fragments	028972T6 (SPLNFET01), 028972T6 (SPLNFET01), 2 452387R6 (TLYMNOT02), 736580R1 (TONSNOT01), 7 936731R1 (CERVNOT01), 3206282H1 (PENCNOT03), (SPLNNOT09), 3742964H1 (THYMNOT08), 4028320H1 4726757H1 (GBLADIT01), 5473562H1 (MCLRUMT01)		1281527H1 (COLNNOT16), 1412985H1 (BRA (PANCNOT04), 2 2470481F6 (THP1NOT03), 2539015F7 (BON (BRSTTUT15), 3694831H1 (PANCNOT19), 3	111466R1 (PITU 2501034H1 (ADR), 008963H1 (H), 723168X19 (9H1 (BSTMNON02 1), 4920466H1	864776T1 (BRAITUTO3), 1911267F6 (CONN (BRAVTXT03), SARA02608F1, SARA02002F1	876677H1 (LUNGAST01), 876677R6 (LUNGAST01),	867305R1 (BRAITUT03), 963058R2 (BRSTTUT03), 17 (UCMCNOT02), 1727927T6 (PROSNOT14), 2326143H1 2326143R6 (OVARNOT02), 3360563H1 (PROSTUT16)	278630241 (BRSTNOT13), 2958321X303D1 (ADRENOT09), (ADRENOT09), 2958321X308D1 (ADRENOT09)	551126H1 (BEPINOT01), 2808373H1 (BLADTUT08), (SMCCNOS01), 3735780H1 (SMCCNOS01), 3735780T 4760604T6 (BRAMNOT01)	039026H1 (HUVENOB01), 159164F1 (ADENINB01),	063159R6 (PLACNOBO1), 260607R6 (HNT2RAT01), (TESTTUT02), 1273069H1 (TESTTUT02), 2867453 3082466H1 (BRAIUNT01), 4796739H1 (LIVRTUT09
Library	HNT2AGT01	COLNFET02	PANCNOT04	ADRETUT05	LUNGTUT11	BRAVTXT03	LUNGAST01	OVARNOT02	BRSTNOT13	SMCCNOS01	HUVENOB01	HNT2RAT01
Clone ID	264114	1455669	2084989	2501034	2745212	4833111	876677	2326143	2786302	3735780	039026	260607
Nucleotide SEQ ID NO:	44	45	46	47	48	49	50	51	52	53	54	55
Protein SEQ ID NO:	1	2	E.	4	S.	9	7	в	6	10	11	12

Table 1 (cont.)

Fragments	1429651F1 (SINTBST01), 1429651H1 (SINTBST01), 1501096F6 (SINTBST01), 1989621T6 (CORPNOT02), SXLA01343V1, SXLA01559V1, SXLA00812V1	2069606F6 (ISLTNOT01), 2069971H1 (ISLTNOT01), 2374634F6 (ISLTNOT01), 2383754F6 (ISLTNOT01), 4171186T6 (SINTNOT21), SXLA01128V1, SXLA01348V1, SXLA01219V1, SXLA00260V1, SXLA00074V1	TTO4), 1657742 9339H1 (COLNN IOTO3), 914940	. 0	883601R1 (PANCNOTO5), 1525902F6 (UCMCL5T01), 1525902X18C1 (UCMCL5T01), 1525902X18C1 (UCMCL5T01), 1527325T6 (UCMCL5T01), 1554770X311D1 (BLADTUT04), 2417265H1 (HNT3AZT01), 2444786F6 (THPINOTO3), 2722462H1 (LUNGTUT10), 4293114H1 (BRABDIR01), 5070268T6 (PANCNOT23), SANA01850F1, SAJA01078R1, SANA02081F1, SAJA01813F1	000573H1 (U937NOT01), 494409F1 (HNT2NOT01), 494409R1 (HNT2NOT01), 2506506F6 (CONUTUT01), 2681059H1 (SINIUCT01), 2744648F6 (BRSTTUT14), 2805590F6 (BLADTUT08), 3770643H1 (BRSTNOT25), 4204278H1 (BRAITUT29), SAEA02093F1	487309R7 (HNT2AGT01), 1361439F1 (LUNGNOF12), 2758310H1 (THP1AZS08), SCFA05584V1, SCFA05940V1, SCFA0516V1, SCFA05135V1	632097R6 (KIDNNOT05), 632097T6 (KIDNNOT05), 2762348H1 (BRSTNOT12), SCCA02837V1, SCCA05356V1, SCCA01377V1, SCCA05963V1, SCCA05364V1, SCCA02307V1, SCCA04327V1, SCCA02009V1	3723F6 (PROSNC UNGFET05), 236 15961H1 (PENCN RABDIR01)	1322651X35 (BLADNOT04), 1322651X36 (BLADNOT04), 3494841H1 (ADRETUT07), 4958978F6 (TLYMNOT05), 5108194H1 (PROSTUS19), 91379009, 91527417	
Library	SINTBST01	ISLTNOT01	COLNNOT11	BONRTUT01	LUNGTUT10	OVARNOT09	THP1AZS08	BRSTNOT12	PENCNOT09	PROSTUS19	BRABDIR01
Clone ID	1429651	2069971	2329339	2540219	2722462	2739264	2758310	2762348	3715961	5108194	5503122
Nucleotide SEQ ID NO:	56	57	58	59	09	61	62	63	64	65	99
Protein SEQ ID NO:	13	14	15	16	17	18	19	20	21	22	23

Fragments	805957R1 (BSTMNOT01), 953622R1 (SCORNON01), 1501080F1 (SINTBST01), 1547381R6 (PROSTUT04), 2081843T6 (UTRSNOT08), 2811524F6 (OVARNOT10), 3212921H1 (BLADNOT08), 3250443H1 (SEMVNOT03), 3269479H1 (BRAINOT20), 3699955F6 (SININOT05), 3700568H1 (SININOT05), 4944050H1 (BRAIFEN05), 5517972H1 (LIVRDIR01)	2859465F6 (SININOTO3), 2859465T6 (SININOTO3), 3555656F6 (LUNGNOT31), 3555656T6 (LUNGNOT31), 4345952H1 (TLYMTXT01), 5593114H1 (COLCDITO3), 5874544H1 (COLTDITO4)	044775H1 (TBLYNOT01), 044775X3 (TBLYNOT01), 455640R1 (KERANOT01), 950702R1 (PANCNOT05), 2418550H1 (HNT3AZT01), 2798917H1 (NPOLNOT01), 2844696H1 (DRGLNOT01), g1718929	699714R6 (SYNORATO3), 831423R1 (PROSTUTO4), 978875R1 (BRSTNOT02), 1350569F1 (LATRTUTO2), 1447681R1 (PLACNOTO2), 3177382F6 (UTRSTUTO4), 3688796H1 (HEAANOTO1), 3929008H1 (KIDNNOT19), g2106455, g2163092	571573F1 (OVARNONO1), 571573R1 (OVARNONO1), 875369H1 (LUNGAST01), 875369R1 (LUNGAST01), 3569021H1 (HEAPNOT01)	1325518H1 (LPARNOTO2), 1325518T6 (LPARNOTO2), 1825553F6 (LSUBNOTO3), SBAA02035F1	1378947T1 (LUNGNOT10), 1453290F1 (PENITUT01), 1459818R1 (COLNFET02), 1967477H1 (BRSTNOT04), 2060987H1 (OVARNOT03), 2455371F6 (ENDANOT01), 2499967F7 (ADRETUT05), 3093056T6 (BRSTNOT19), 3213366H1 (BLADNOT08), 4934158H1 (BRSTTUT20), SBYA01942U1	2172064CT1 (ENDCNOT03), 2172064H1 (ENDCNOT03), SBLA01269F1 2219267F6 (LUNGNOT18), 2219267H1 (LUNGNOT18), 3117478T6 (LUNGTUT13), 3126288T6 (LUNGTUT12), 3558495H1 (LUNGNOT31)	MMLRIDTO1), 469862R1 (MMLRIDTO1), 150), 2191933H1 (THYRTUTO3)	1326594F1 (LPARNOTO2), 2256143H1 (OVARTUTO1), 2278689R6 (PROSNONO1), 2528425H1 (GBLANOTO2), 2660038H1 (LUNGTUTO9), 2449964H1 (UTRSNONO3), 5099879H1 (PROSTUS20), 41886680, 4783969
Library	LIVRDIR01	COLCDIT03	TBLYNOT01	KIDNNOT01	LUNGAST01	LPARNOT02	OVARNOT03	ENDCNOT03 LUNGNOT18	NGANNOT01	LUNGTUT09
Clone ID	5517972	5593114	044775	116588	875369	1325518	2060987	2172064 2219267	2308629	2660038
Nucleotide SEQ ID NO:	67	89	69	70	71	72	73	74	76	77
Protein SEQ ID NO:	24	25	26	27	28	29	30	31	33	34

- 67					 	_		 		_		
	Fragments	259200X12 (HNT2RAT01), 1266477F1 (BRAINOT09), 2383364F6 (ISLTNOT01), 2670745H1 (ESOGTUT02), 3181526H1 (TLYJNOT01)	607375K6 (BRSTTUT01), 1728626X15C1 (PROSNOT14), 1751773F6 (LIVRTUT01), 1751994T6 (LIVRTUT01), 1796032X14C1 (PROSTUT05), 2010172H1 (TESTNOT03), 2676443H1 (KIDNNOT19)	063264H1 (PLACNOB01), 434468T6 (THYRNOT01), 487721H1 (HNT2AGT01), 907796R2 (COLNNOT09), 1212556R7 (BRSTTUT01), 1251889H1 (LUNGFET03), 16533706 (PROSTUT08), 1653370x309D1 (PROSTUT08), 10376785 (THYRNOTOR), 236576781	3438320H1 (PENCNOT06), 3501438H1 (PROSTUT13), 3745542H1 (THYMNOT08), 3751060H1 (UTRSNOT18), 4979750F6 (HELATXT04),	SADA00043F1, SADA00087F1	1634141F6 (COLNNOT19), 1692115X12C1 (PROSTUT10), 1731310F6	1318962H1 (BLADNOTO4), 1520864F1 (BLADTUTO4), 1684381F6 (PROSNOT15), 2055747R6 (BEPINOTO1), 4378816H1 (LUNGNOT37)	4797137F6 (LIVRTUT09), 4797137H1 (LIVRTUT09), 4797137T6	(LIVRTUT09)	5470806H1 (MCLRUNT01), 5470806T6 (MCLRUNT01)	5473242F6 (MCLRUNT01), 5473242T6 (MCLRUNT01)
	Library	ESOGTUT02	KIDNNOT19	TLYJINTOI	PENCNOT06		UTRSTUT05	LUNGNOT37	LIVRTUT09		MCLRUNT01	MCLRUNT01
	Clone ID	2670745	2676443	3295764	3438320		3986488	4378816	4797137		5470806	5473242
	Nucleotide SEQ ID NO:	78	62	80	81		82	83	84		85	86
	Protein SEQ ID NO:	35	36	37	38		39	40	41		42	43

Table 2

Analytical Methods & Databases	MOTIFS BLIMPS-PFAM BLIMPS-PRINTS BLAST-GenBank BLAST-PRODOM BLAST-DOMO	MOTIFS BLAST-GenBank BLAST-PRODOM SPSCan HMMER	MOTIFS BLAST-GenBank SPScan	MOTIFS BLAST-GenBank SPScan HMMER	BLAST-GenBank MOTIFS	BLAST-GenBank HWMER-PFAM MOTIFS ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO SPECAT
Homologous Sequences	Ring canal protein [Drosophila melanogaster] g577276	Multi-drug resistance- associated protein (MRP)-like protein. 1 (MLP-1) [Rattus norvegicus]	Tricarboxylate carrier [Rattus sp.] g545998	Weak similarity with honeybee ATP synthase A chain [Caenorhabditis elegans] q3878801	Cu2+-transporting ATPase homolog [Arabidopsis thaliana] g2464854	1 # a
Signature Sequences, Motifs, and Domains	BTB domain: C44-F56 POZ domain: N10-Q211 Kelch repeat signature: B379-G392, T398-V412, L438-M452, T498-A512 Ring canal protein repeat: E122-P254	Signal peptide: M1-G36 Transmembrane region: S25-W45 MRP(2) MRP(1) repeat: C30-V74	Signal peptide: M1-N52	Signal peptide: M1-C30 Transmembrane region: L233-F252	Leucine zipper: L284-L305	Mitochondrial energy transfer proteins: G5-L266 Signal peptide: M1-G17
Potential Glycosylation Sites	N97 N333	N15	N103 N127 N135 N138			
Amino Potential Acid Phosphorylation sidues Sites	S521 S2 T3 S16 S99 S138 S144 T193 T264 T404 S448 S589 S151 T229 T337 T457 S562 S568	T17	T334 T33 S137 T146 S291 S311 T346	S234 T126 T169 Y141	S99 S125 S192 T277 S307 S309 T110 Y212	S96 T198 S215 T29 S121 S164 S170
Amino Acid Residues	623	66	374	271	323	274
SEQ ID NO:	П	2	3	7	2	9

Analytical Methods & Databases	MOTIFS BLAST-GenBank SPScan HWMER BLIMPS-BLOCKS BLAST-DOMO BLAST-PRODOM	MOTIFS BLAST-GenBank	MOTIFS BLAST-GenBank BLAST-DOMO	MOTIFS BLAST-GenBank HMWER-PFAM BLIMPS-PRINTS	BLAST-GenBank MOTIFS	MOTIFS SPScan HMMER ProfileScan
Homologous Sequences	Stomatin [Homo MC sapiens] g1161562 BI SI	K* channel MC modulatory factor BI DEBT-91 [Mus musculus] 94838557	ABC2 transporter MO [Mus musculus] BL g495259 BL	Similar to human MO ADP/ATP carrier BL protein [C. HM elegans] g3879938 BL	Mitochondrial BL import protein MO Tim9p [Saccharomyces cerevisiae]	
Signature Sequences, Motifs, and Domains	Signal peptide: M1-T42 Transmembrane domain: W29-I54 Band 7 protein family: C50-V62, K90-E224 Membrane stomatin: E14-N283		ABC transporter family: RR79-K177 ATP/GTP-binding site motif A (P-loop): G102-S109	Mitochondrial carrier protein signature: RE117-1297 FGraves Disease carrier protein: P137-T157, L259-S279		Signal peptide: M1-G24 Transmembrane domain: G35-F57 Sodium neurotransmitter symporter signature:
Potential Glycosylation Sites	N226 N261	N218 N253 N259	N87	N287		
Potential Phosphorylation Sites	S6 T113 T173 T147 S230 T258	S2 S35 T57 S92 T104 S191 S302 S334 S335 S336 T43 T250 T255 T304 S311 S370 Y65	T160 S17 T71 S77 T78 S111 S134 S142	S17 S114 T136 S16	T37 T47 T60 S64	T108 T84
Amino Acid Residues	291	381	190	297	88	115
SEQ ID NO:	7	œ	6	10	11	12

Table 2 (cont.)

Analytical Methods & Databases	BLAST-GenBank NOTIFS HNMER HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-DOMO	BLAST-GenBank MOTIFS HWMER BLIMPS-PRODOM BLAST-PRODOM BLAST-DOMO	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO	BLAST-GenBank MOTIFS SPScan HMMER-PFAM ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS
Homologous Sequences	Sodium-glucose cotransporter [Oryctolagus cuniculus] g473969	Zinc transporter ZnT-2 [Rattus norvegicus] g1256378	Ring canal protein [Drosophila melanogaster] g577276	Carrier protein (c1) [Caenorhabditis elegans] g472902
Signature Sequences, Motifs, and Domains	Transmembrane domains: 129-V48, L103-1121, L177-G196, 1210-M229, L417-W435, F481-Y501, Y521-W541 Sodium symporter family domain: Y58-G487 Sodium:solute symporter signature: Y35-G89, M111-R140, L173-G227, P460-G469	Transmembrane domains: 192-L112, 1201-K219 Zinc transporter signature: A28-V142, D199-E303 Cation transporter domain: S48-L74	Kelch repeat motifs: C299-N349; F350-R399 Y400-G446 BTB domain: F50-L117 POZ domain: Y27-E215	Signal peptide: M1-S17 Mitochondrial carrier proteins domain: C4-189 Mitochondrial carrier proteins signature sequence: V6-G19, G19-A33, G63-E83
Potential Glycosylation Sites	N243 N247 N301 N601	N162 N234	N24 N279	
Potential Phosphorylation Sites	T54 T50 S99 T127 S413 T558 S645 T654 T47 S242 T602 T611 Y501	T84 S304 T11 S75 S80 S164 Y20	S111 S145 S183 S233 T26 T185 S202 T243	T22 Y37
Amino Acid Residues	675	320	462	86
SEQ ID NO:	13	14	15	16

Analytical Methods & Databases	BLAST-GenBank MOTIFS SPScan HMMER HMMER-PFAM BLIMPS-PRINTS	BLAST-GenBank MOTIFS SPScan	BLAST-GenBank MOTIFS BLIMPS-PRODOM BLAST-PRODOM	BLAST-GenBank MOTIFS HWMER BLIMPS-PRODOM BLAST-PRODOM BLAST-DOMO	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-DOMO
A E O	BLAST- MOTIFS SPScan HWMER HWMER- BLIMPS-	BLAST- MOTIFS SPScan	BLAST-OMOTIFS BLIMPS BLAST-		BLAST- MOTIFS HMMER- BLIMPS Profil- BLAST- BLAST-
Homologous Sequences	Voltage-gated calcium channel [Rattus norvegicus] g4586963	Nucleoporin p54 [Rattus norvegicus] g1537070	ABC transporter [Methanobacterium thermo.] g2622773	Vacuolar H+/ATPase subunit [Rattus norvegicus] g206430	Mitochondrial uncoupling protein UCP-4 (Homo sapiens] g4324701
Signature Sequences, Motifs, and Domains	Signal peptide: M1-A61 Transmembrane domains: L39-L56, I167-F186, C229-F252, G438-L455, M492-F509, L598-I618 Ion transport proteins signature: F85-V251, L369-I618	Signal peptide: M1-G26	ABC1 precursor signature: N153-Q162, F210-A229, G234-1254, V312-G332, T366-V378	Transmembrane domains: Y451-D469, M544-F562, F577-F597, G775-M797 Vacuolar ion transport subunit signature: M10-F831	Mitochondrial carrier proteins domain: Y31-S248 Mitochondrial energy transfer proteins signature sequence: 162-086, 1110-G122
Potential Glycosylation Sites	N548 N627 N548 N627	N220 N250 N364 N496		N368 N490 N624	
Potential Phosphorylation Sites	S55 S196 T254 S307 S327 T491 T534 T550 T571 S635 S648 S677 T696 S283 S291 T314 S629 S701 Y556	T200 S183 T232 T284 T349 T150 T252 S253 S319 S383 Y454	S460 S104 T178 S320 S321 T498 T531 Y365	T98 S120 S203 T214 T276 S388 T438 T700 T838 T167 T179 S280 T370 S435 S531 S539 S666 S693 S830	S50 T139 T152 T177 S202 T143 Y55
Amino Acid Residues	748	507	592	841	253
SEQ ID NO:	17	18	19	20	21

Analytical Methods & Databases	BLAST-GenBank MOTIFS SPSCan HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PFAM BLAST-PRODOM BLAST-DOMO	BLAST-GenBank MOTIFS HWMER HWMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-PRODOM	BLAST-GenBank MOTIFS	MOTIFS HWMER-PFAM BLAST-PRODOM BLAST-DOMO BLAST-GenBank
Homologous Sequences	Grave's disease carrier protein [Bos taurus] g387	Voltage-dependent calcium channel beta-4 subunit [Homo sapiens] g2058727	Breast cancer resistance protein (multidrug transporter) (Homo sapiens] g4038352	Cation transport protein [E. coli] g495778	Similar to carrier protein C2 [C. elegans] g3879669
Signature Sequences, Motifs, and Domains	Signal peptide: M1-A47 Mitochondrial carrier proteins domain: Q32-G220 Mitochondrial carrier proteins signature sequence: S36-T49, T49-V63, G92-E112, T144-T162, Y187-F205	Dihydroxipyridine- sensitive L-type calcium channel signature: Y2-A47, I49-V77, A83-N100, R106-E131 SH3 domain: V59-R122	Transmembrane domains: I396-K417, Y494-S522, T538-V556 ABC transporters domain: P73-G262 ABC transporter family signature sequence: I78-L89, V186-D217		Mitochondrial energy transfer proteins signatures: P89-L97, M1-E41, M73-L152 Mitochondrial carrier
Potential Glycosylation Sites		N66 N145	N338 N418 N557 N596	N27	
Potential Phosphorylation Sites	S69 S26 S109 T162 S178 S25 S64 S65 T210 S219	S26 S31 S149 S164 T22 T157	T194 S195 S232 T362 S655 S4 S88 T135 T153 S187 T214 S322 T345 S353 S443 T609 S261 S381	T51 S29 T100 S138 S151 Y78	S54 S42 S62 T78 Y104
es les	229	170	655	184	154
SEQ ID NO:	22	23	24	25	26

Analytical Methods & Databases	MOTIFS HWMER BLAST-GenBank	MOTIFS BLAST-DOMO BLAST-GenBank	MOTIFS HMMER SPScan ProfileScan BLAST-GenBank	MOTIFS BLIMPS-PRINTS BLAST-DOMO BLAST-GenBank	MOTIFS SPSCan HWMER BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PFAM ProfileScan BLAST-PRODOM BLAST-DOMO BLAST-GenBank
Homologous Sequences	Multidrug efflux transporter [Bacillus subtilis] g2635104	ARL-6 interacting protein-4 [Mus musculus] g4927204	Surface antigen [Trypanosoma cruzi] g161956	NY-REN-45 antigen (similar to potassium channel protein) (Homo sapiens] g5360115	Gap junction protein (similar to connexin) [Homo sapiens] g3006230
Signature Sequences, Motifs, and Domains	Transmembrane domains: C91-L111, L237-I257, I305-M332, M332-L354, L391-V408, I186-A204	Nucleic acid-binding protein E5.1 domain: S6-K128	Signal peptide: M1-R19 or M1-K15 Caseins alpha/beta signature: M1-N39	Potassium channel signature: A62-T81 Potassium channel integral membrane protein domain: S13-D117	Signal cleavage: M1-G45 Connexin domains: M1-V99. V20-Y44 Connexin signatures: L33-V86, L152-F205, F51-P73, S76-L96, L133-Y159, C169-T189, I190-L218 Gap junction protein connexin transmembrane regions: M1-S130
Potential Glycosylation Sites	NSO N423	N35		N343 N570 N638 N703	N181
Potential Phosphorylation Sites	S170 T5 T51 T265 T300 S425	S10 S47 T72 S28 S96 S148 T173 T222 S6 S21 T32 T61 T192	T66 S194 T200	S31 T6 T55 T263 T328 T546 T580 T594 S662 S673 T32 S50 S231 T244 T306 T385 S439 S476 S533 S553 S624	T18 T245 T206
Amino Acid Residues	438	237	219	707	279
SEQ ID NO:	27	28	29	30	31

WO 00/78953

PCT/US00/16668

Table 2 (cont.)

Analytical Methods & Databases	MOTIFS HMMER SPScan BLAST-GenBank	MOTIFS HWMER-PFAM BLAST-DOMO BLAST-PRODOM ProfileScan BLAST-GenBank	MOTIFS HWMER-PFAM ProfileScan BLAST-PRODOM BLAST-DOMO BLAST-GenBank	MOTIFS HIMER BLAST-PFAM BLAST-GenBank
Homologous Sequences	mBOCT (potent organic cation transporter) [Mus musculus] 94589468	Mitochondrial solute carrier [Onchocerca volvulus] g1518458	YKL522=mitochondria ADP/ATP carrier protein homolog Saccharomyces cerevisiae g254449	Similarity to Human host cell factor C1 [Homo sapiens] q3875291
Signature Sequences, Motifs, and Domains	Signal peptide: M1-A35 or M1-A14 Transmembrane domain: F83-L102	Mitochondrial energy transfer proteins signatures: M1-G147, P17, P115, N185-K280, A101-Q181, Y184-I278 Mitochondrial carrier protein domains: M1-E176, N185-K280 Mitochondrial transmembrane transport protein regions: P17-R182, P180-I278	Mitochondrial energy transfer proteins signatures: P19-M27, D2-153, L209-L295 Mitochondrial carrier protein domain: D2-Y295 Transport protein domain: P122-Y295	Kelch motifs: H191-G249, E250-D301
Potential Glycosylation Sites		N60		N96 N372
Potential Phosphorylation Sites	S114	т83 т205 S269 т279	S189 S195 S204 T257	S34 S207 T221 S312 T40 S53 T112 T117 T277 S337
Amino Acid Residues	154	289	300	382
SEQ ID NO:	32	33	34	35

us Analytical es Methods & Databases	1 MOTIFS E HMMER-PFAM LUS BLIMPS-PRINTS BLAST-DOMO BLAST-GenBank	HMMER HMMER HMMER BLAST-PRODOM BLAST-DOMO SPScan BLAST-GenBank	MOTIFS HAWER-PFAM anthus BLAST-GenBank 729	otein MOTIFS HMMER gar BLAST-DOMO [Homo BLAST-GenBank
Homologous Sequences	Mitochondrial dicarboxylate carrier [Rattus norvegicus] g3646426	Reduced folate carrier [Homo sapiens] g1041934	cytochrome b5 containing fusion protein [Helianthus annus] g1040729 P=1.2e-07	Sqv-7-like protein (similar to nucleotide-sugar transporters) [Hom
Signature Sequences, Motifs, and Domains	Mitochondrial energy transfer proteins signatures: P26-L34, P219-L227, L97-G193, W10-V89, D197-F281, P96-Y194 Mitochondrial carrier protein domain: A5-F281 Mitochondrial brown fat region: Y82-Q94, V151-S168, Y194-C212	Transmembrane domains: M114-T137, M364-M380, Y390-A413, A421-D444, F456-V478 Folate transporter domains: W30-H218, I253-K484	Heme-binding domain in cytochrome b5: Y19-G98 Cytochrome b5 family domain: H28-P75	Transmembrane domains: L85-N105, F180-Y200 Intermembrane space domain:
Potential Glycosylation Sites		N63 N314 N414		N214
oten phor Sit	T36 T118 S180 S230 T84 S168 T244	T65 T135 S147 T360 S8 T22 S45 S291	T21 S124 T145 S158 T190 T95 S132 S137 T177	T63 S158 T48
Amino Acid Residues	287	497	228	273
SEQ ID NO:	9 က	37	38	39

Analytical Methods & Databases	MOTIFS HWMER ProfileScan BLAST-DOMO BLAST-GenBank	MOTIFS HWMER BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO BLAST-GenBank	MOTIFS HMMER ProfileScan BLAST-PRODOM BLAST-DOMO BLAST-GenBank BLIMPS-BLOCKS	MOTIFS HWMER-PFAM BLAST-PRODOM BLAST-DOMO BLIMPS-BLOCKS BLIMPS-PRINTS
Homologous Sequences	C-8,7 sterol isomerase, aSI1 [Arabidopsis thaliana] g2772934	Myelin protein zero (MPZ) [Homo sapiens] g2160399	Transthyretin precursor [Sus scrofa] g1009702	III beta-3 globin [Rattus norvegicus] g395943
Signature Sequences, Motifs, and Domains	Signal peptide: M1-G29 or M1-A27 Emopamil binding protein: G37-S187, L15-K203 Transmembrane domain: Y164-L183	Transmembrane domain: F15-I34, M155-V174 Channel myelin protein: L18-M181 Sodium channel beta-2 subunit precursor: F15-E210 Immunoglobulin domain: I34-V136	Signal peptide: M1-G23 or M1-A20 Transthyretin signature: S28-S132 Transthyretin domain: G21-Q146	Globin domain: V2-H147 Heme oxygen transport protein domain: L32-H147
Potential Glycosylation Sites	N158	N123	N118	
Potential Phosphorylation Sites	S187 S201	S192 S200 S56 T95 T146 S199 T207 S229 T53 T61 T69 T119 T148 Y70	T79 T116 S3 S66 Y89 Y98	т5 S88 т39
Amino Acid Residues	206	235	147	147
SEQ ID NO:	40	41	42	43

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
44	1567-1611 2107-2151	Gastrointestinal (0.203) Hematopoietic/Immune (0.188) Nervous (0.156)	Cell Proliferation and Cancer (0.547) Inflammation (0.422)	PBLUESCRIPT
45	1-92 351-434	Endocrine (0.333) Developmental (0.167) Gastrointestinal (0.167) Musculoskeletal (0.167) Reproductive (0.167)	Cell Proliferation and Cancer (0.833) Inflammation (0.167)	pINCY
46	920-964 1352-1396	Reproductive (0.304) Gastrointestinal (0.174) Cardiovascular (0.130) Hematopoietic/Immune (0.130) Nervous (0.130)	Cell Proliferation and Cancer (0.478) Inflammation (0.391)	PSPORT1
47	1-80 768-848	Nervous (0.273) Reproductive (0.273) Gastrointestinal (0.127) Hematopoietic/Immune (0.127)	Cell Proliferation and Cancer (0.564) Inflammation (0.400)	pincy
48	111-194 687-758	Reproductive (0.221) Nervous (0.185) Gastrointestinal (0.124)	Cell Proliferation and Cancer (0.552) Inflammation (0.343)	pINCY
49	1-97	Nervous (0.234) Hematopoietic/Immune (0.191) Gastrointestinal (0.149)	Cell Proliferation and Cancer (0.617) Inflammation (0.340)	pincy
50	218-262	Cardiovascular (1.000)	Cancer (0.333) Inflammation/Trauma (0.333) Cell Proliferation (0.333)	PSPORT1
51	811-855	Hematopoietic/Immune (0.180) Gastrointestinal (0.146) Reproductive (0.281)	Cancer (0.393) Inflammation/Trauma (0.515) Cell Proliferation (0.146)	PSPORT1
52	595-639	Gastrointestinal (0.286) Reproductive (0.714)	Cancer (0.429) Inflammation/Trauma (0.429)	pincy
53	96-140	Cardiovascular (0.167) Hematopoietic/Immune (0.167) Nervous (0.250) Reproductive (0.167)	Cancer (0.250) Inflammation/Trauma (0.167) Cell Proliferation (0.167)	pINCY

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
54	507-551	Reproductive (0.323) Gastrointestinal (0.154) Nervous (0.123)	Cancer (0.446) Inflammation/Trauma (0.308) Cell Proliferation (0.185)	PBLUESCRIPT
55	455-499	Urologic (0.333) Nervous (0.222) Reproductive (0.222)	Cancer (0.667) Cell Proliferation (0.333)	PBLUESCRIPT
56	1835-1879	Nervous (0.625) Gastrointestinal (0.375)	Inflammation/Trauma (0.375) Cancer (0.250) Neurological (0.250)	pincy
57	811-855	Gastrointestinal (1.000)	Inflammation/Trauma (0.667)	PINCY
58	390-434	Reproductive (0.320) Nervous (0.240) Urologic (0.120)	Cancer (0.520) Inflammation/Trauma (0.240) Cell Proliferation (0.160)	PSPORT1
59	413-457	Gastrointestinal (0.333) Musculoskeletal (0.333) Nervous (0.333)	Cancer (0.333) Neurological (0.333)	pINCY
09	2021-2084	Nervous (0.197) Gastrointestinal (0.184) Reproductive (0.184)	Cancer (0.461) Inflammation/Trauma (0.316) Cell Proliferation (0.118)	DINCY
19	62-109	Nervous (0.226) Reproductive (0.208) Cardiovascular (0.113) Gastrointestinal (0.113)	Cancer (0.528) Inflammation/Trauma (0.301) Cell Proliferation (0.208)	pINCY
62	379-423 1867-1911	Reproductive (0.282) Gastrointestinal (0.205) Nervous (0.154)	Cancer (0.538) Inflammation/Trauma (0.282) Cell Proliferation (0.103)	PSPORT1
63	362-406 1193-1237	Urologic (0.500) Reproductive (0.333) Cardiovascular (0.167)	Cancer (0.667) Inflammation/Trauma (0.333)	pINCY
64	394-438	Nervous (0.294) Reproductive (0.265) Cardovascular (0.118)	Cancer (0.382) Inflammation/Trauma (0.235) Cell Proliferation (0.118)	pincy

Vector	pincy	pINCY	pincy	pincy	PBLUESCRIPT	PBLUESCRIPT	PSPORT1	PINCY	PSPORT1	pincy	pincy	PSPORT1
Disease or Condition (Fraction of Total)	Inflammation/Trauma (0.500) Cancer (0.400)	Neurological (1.000)	Cancer (0.500) Inflammation/Trauma (0.294) Cell Proliferation (0.118)	Inflammation/Trauma (0.546) Cell Proliferation (0.182)	Cancer (0.250) Cell Proliferation (0.375) Inflammation/Trauma (0.416)	Cancer (0.373) Inflammation/Trauma (0.382) Cell Proliferation (0.176)	Cancer (0.438) Inflammation/Trauma (0.314) Cell Proliferation (0.176)	Cancer (1.000)	Cancer (0.459) Inflammation/Trauma (0.379) Cell Proliferation (0.203)	Cancer (0.250) Cell Proliferation (0.250) Inflammation/Trauma (0.500)	Cancer (0.571) Cell Proliferation (0.286) Inflammation (0.143)	Cancer (0.494) Inflammation (0.215) Cell Proliferation (0.127)
Tissue Expression (Fraction of Total)	Reproductive (0.300) Endocrine (0.200) Gastrointestinal (0.200) Hematopoietic/Immune (0.200)	Nervous (1.000)	Reproductive (0.324) Nervous (0.265) Gastrointestinal (0.235)	Hematopoietic/Immune (0.455) Gastrointestinal (0.182) Nervous (0.182)	Nervous (0.292) Gastrointestinal (0.208) Hematopoietic/Immune (0.125)	Reproductive (0.206) Hematopoietic/Immune (0.186) Cardiovascular (0.127)	Reproductive (0.275) Nervous (0.163) Gastrointestinal (0.137)	Gastrointestinal (1,000)	lw ⊨	Nervous (0.750) Dermatologic (0.250)	Cardiovascular (0.714) Developmental (0.143) Hematopoietic/Immune (0.143)	Reproductive (0.253) Nervous (0.241) Gastrointestinal (0.127) Hematopoietic (0.127)
Selected Fragments	768-812	77-121	1999-2043	561-605	679-729	95-366 1078-1185	33-152	81-779	719-817 1202-1414	1-848	1-478	1-134
Nucleotide SEQ ID NO:	65	99	29	89	69	70	71	72	73	74	75	76

Vector	pincy	PINCY	pincy	PINCY	PINCY	pincy	pincy	pincy	PINCY	pINCY
Disease or Condition (Fraction of Total)	Cancer (0.467) Inflammation/Trauma (0.467)	Cancer (0.478) Inflammation/Trauma (0.292) Cell Proliferation (0.191)	Cancer (0.750) Cell Proliferation (0.125) Inflammation/Trauma (0.167)	Cancer (0.490) Inflammation/Trauma (0.286) Cell Proliferation (0.224)	Cancer (0.500) Cell Proliferation (0.333) Inflammation (0.167)	Inflammation/Trauma (0.461) Cancer (0.308) Cell Proliferation (0.205)	Cancer (0.564) Cell Proliferation (0.256) Inflammation/Trauma (0.180)	Cancer (0.500) Inflammation (0.500)	Cell Proliferation (1.000)	Cell Proliferation (1.000)
Tissue Expression (Fraction of Total)	Reproductive (0.467) Cardiovascular (0.133) Gastrointestinal (0.133)	Reproductive (0.230) Nervous (0.225) Gastrointestinal (0.124)	Reproductive (0.417) Gastrointestinal (0.292) Urologic (0.125)	Reproductive (0.245) Nervous (0.143) Developmental (0.122)	Reproductive (0.667) Cardiovascular (0.167) Nervous (0.167)	Gastrointestinal (0.282) Hematopoietic/Immune (0.205) Reproductive (0.205)	Nervous (0.179) Reproductive (0.179) Gastrointestinal (0.128)	Gastrointestinal (0.500) Hematopoietic/Immune (0.500)	Hematopoietic/Immune (1.000)	Developmental (1.000)
Selected Fragments	510-719 960-1100	180-293	192-653 795-935	139-1044	233-916	1-153 760-816	57-299	1-707	451-594	8-124 161-187 407-472
Nucleotide SEQ ID NO:	77	78	79	80	81	82	83	84	85	98

Table 4

SEQ Library ID NO: 44 HNT2AGT01 45 COLNFET02 47 ADRETUT05 48 LUNGTUT11 50 LUNGAST01 51 OVARNOT02 52 BRSTNOT13

SEQ ID NO:	Library	Library Comment
53	SMCCNOS01	Library was constructed using 7.56 X 10e6 clones from a coronary artery smooth muscle cell library and was subjected to two rounds of subtraction hybridization for 48 hours with 6.12 X 10e6 clones from a control coronary artery smooth muscle cell library. The starting library for subtraction was constructed using RNA isolated from coronary artery smooth muscle cells removed from a 3-year-old Caucasian male. The cells were treated with TNF alpha & IL-1 beta 10ng/ml each for 20 hours. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated coronary artery smooth muscle cells from the same donor.
54	HUVENOB01	Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells.
52	HNT2RAT01	Library was constructed at Stratagene (STR937231), using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated with retinoic acid for 24 hours.
56	SINTBST01	Library was constructed using RNA isolated from ileum tissue obtained from an 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
57	ISLTNOT01	Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.
58	COLNNOT11	Library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy.
59	BONRTUT01	Library was constructed using RNA isolated from rib tumor tissue removed from a 16-year-old Caucasian male during a rib osteotomy and a wedge resection of the lung. Pathology indicated metastatic grade 3 (of 4) osteosarcoma, forming a mass involving the chest wall.
09	LUNGTUT10	Library was constructed using RNA isolated from lung tumor tissue removed from the left upper lobe of a 65-year-old Caucasian female during a segmental lung resection. Pathology indicated metastatic grade 2 myxoid liposarcoma and metastatic grade 4 liposarcoma. Patient history included soft tissue cancer, breast cancer, and secondary lung cancer.
61	OVARNOT09	Library was constructed using RNA isolated from ovarian tissue removed from a 28-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Pathology indicated multiple follicular cysts ranging in size from 0.4 to 1.5 cm in the right and left ovaries, chronic cervicitis and squamous metaplasia of the cervix, and endometrium in weakly proliferative phase. Family history included benign hypertension, hyperlipidemia, and atherosclerotic coronary artery disease.

Library THP1AZS08 BRSTNOT12 PENCNOT09 PROSTUS19 BRABDIR01	Library was constructed using 5.76 million clones from a 5-aza-2'-deoxycytidine (AZ) treated THP-1 promonocyte cell line library. Starting RNA was made from THP-1 promonocyte cells treated THP-1 promonocyte cells in micromolar AZ. 5.76 million clones from the promonocyte cells treated THP-1 promonocyte cells from the untreated THP-1 cell library were then subjected to two rounds of subtractive hybridization conditions were based on the methodolegies of Swaroop et al. [1991] Nucleic Acids Res. 19:1954, and Bonaldo et al. (1996) Genome Research 6:791. THP-1 (ATC TIB 202) is a human promonocyte cell line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref. Int. J. Cancer [1980) 26:1711. Library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mamoplasty. Pathology indicated atherosclerotic coronary artery disease. Family history included benign hypertension and atherosclerotic coronary artery disease. Family history included benign hypertension and atherosclerotic coronary artery disease. Family history included benign mypertension and prostate library. The starting library for subtraction was constructed using RNA isolated from penies right corpora cavernosa tissue from a bubjected to two rounds of subtraction hypridization with 2.36 million clones from a prostate prostate using RNA isolated from penies right corpora experies a subjected to two rounds of subtraction hypridization with a subjected to two rounds of subtraction hypridization with 2.36 million clones from a subjected to two rounds of subtraction as 59-year-old Caucasian male during a radical prostate tumor tissue removed from a 59-year-old Caucasian male during a radical prostate tumor tissue removed from a 59-year-old caucasian male with elevated prostate specific antigen (ESA). Patient history included benigh hypertlapidemia, and rheumatoid arthritis. The patient presented with reduced using RNA isolated from a cerbevo
LIVRDIR01	Library was constructed using RNA isolated from diseased liver tissue removed from a 63-year-old Caucasian female during a liver transplant. Patient history included primary

SEQ ID	Library	Library Comment
89	COLCDIT03	Library was constructed using RNA isolated from diseased colon polyp tissue removed from the cecum of a 67-year-old female. Pathology indicated a benign cecum polyp. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma that arose in tubulovillous adenoma forming a fungating mass in the cecum.
69	TBLYNOT01	
70	KIDNNOT01	as constr female, and tobe
71	LUNGAST01	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma.
72	LPARNOT02	Library was constructed using RNA isolated from tissue obtained from the left parotid (salivary) gland of a 70-year-old male with parotid cancer.
73	OVARNOT 03	Library was constructed using RNA isolated from ovarian tissue removed from a 43-year-old Caucasian female during a bilateral salpingo-oophorectomy. Pathology for the associated tumor tissue indicated grade 2 mucinous cyst adenocarcinoma. The patient presented with stress incontinence, Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, cerebrovascular disease, breast cancer, and uterine cancer.
74	ENDCNOT03	Library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
75	LUNGNOT18	
92	NGANNOT01	Library was constructed using RNA isolated from tumorous neuroganglion tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma. Family history included asthma.
77	LUNGTUT09	sia Sia

ESOGTUTO2 Library was control of seaso, type hypertension, control of control of a male based of the penis of the penis of the penis of the penis of the endocervice hypertension, hippertension, hypertension and prostate of a 3-year-old a 1-penis of the penis of the penis of the endocervice hypertension of the endocervi	
ESOGTUTO2 Library was of year-old Cauce pyloromyotomy adenocarcinon disease, type hypertension, hypertension, caucasian malassociated tuthe left kidr neoplasm of caucasian malassociated tuthe left kidr neoplasm of cand prostate library was of a male. Palencholof Library was of a male. Palencholof Library was of a line library was of a library was of the endocervition of the endocervi	Library Comment
KIDNNOT19 Library was c Caucasian mal associated tu the left kidr neoplasm of c myocardial ir and prostate of a male. Pa line library was c of a male. Pa line library was c a 3-year-old 4 urothelial of the penis 4 urothelial of the penis 4 urothelial of the penis 4 urothelial of the endocervi LUNGNOT37 Library was c old Caucasian cytomegalovir LIVRTUT09 Library was c old Caucasian cytomegalovir Library was c old Caucasian cytomegalovir is a derivati year-old Caucasian cytomegalovir is a derivati	nstructed using sian male during and regional ly in the esophagu II diabetes, chrand cerebrovascu
TLYJINT01 Library was of a male. Pa line library line library a 3-year-old 4 urothelial of the penis 4 urothelial of the penis 4 urothelial of the penis byear-old Cauc Pathology ind fragments of the endocervi LUNGNOT37 Library was of old Caucasian cytomegalovi LIVRTUT09 Library was of cold Caucasian cytomegalovi LIVRTUT09 Library was of cold Caucasian cytomegalovi is a derivati	nstructed using RNA isolated from kidney tissue removed a 65-year-old during an exploratory laparotomy and nephroureterectomy. Pathology for the or tissue indicated a grade 1 renal cell carcinoma within the upper pole of y. Patient history included malignant melanoma of the abdominal skin, benign lon, cerebrovascular disease, and umbilical hernia. Family history included arction, atherosclerotic coronary artery disease, cerebrovascular disease, ancer.
PENCNOT06 Library was cand a 3-year-old 4 urothelial of the penis 4 urothelial of the penis 4 urothelial LURSTUT05 Library was condenoted by the endocervic the endocervic the endocervic condenoted by LUNGNOT37 Library was condenoted by Concept condenoted by Concep	nstructed using RNA isolated from a Jurkat cell line derived from the T cells ient history included acute T-cell leukemia. This is an uninduced Jurkat cell rom the same donor.
UTRSTUTU5 Library was constructed using RNA isolated year-old Caucasian female during a vaginal Pathology indicated uterine leiomyoma. The fragments of endometrial polyps. Benign end the endocervix. Patient history included a LUNGNOT37 Library was constructed using polyA RNA isold Caucasian female who died from a closed cytomegalovirus. LIVRTUT09 Library was constructed using RNA isolated is a derivative of Hep G2, a cell line deriyear-old Caucasian male.	constructed using RNA isolated from penis corpora cavernosa tissue removed from Black male. Pathology for the associated tumor tissue indicated invasive grade carcinoma forming a soft tissue scrotal mass that invaded the cavernous body and encased both testicles. Right inguinal lymph node showed metastatic grade carcinoma, with extranodal invasion.
LUNGNOT37 Library was constructed using polyA RNA iso- old Caucasian female who died from a closed cytomegalovirus. LIVRTUT09 Library was constructed using RNA isolated is a derivative of Hep G2, a cell line deriv- year-old Caucasian male.	nstructed using RNA isolated from uterine tumor tissue removed from a 41-sian female during a vaginal hysterectomy with dilation and curettage. cated uterine leiomyoma. The endometrium was secretory and contained ndometrial polyps. Benign endo- and ectocervical mucosa were identified in Patient history included a ventral hernia and a benign ovarian neoplasm.
LIVRTUT09 Library was constructed using is a derivative of Hep G2, a cyear-old Caucasian male.	poly? from
	tructed using RNA of Hep G2, a cell an male.
MCLRUNT01 Library was	structed using RNA isolated from untreated peripheral blood mononuclear from buffy coat, removed from a 60-year-old male.
86 MCLRUNT01 Library was constructed using RNA isolated from Lissue obtained from buffy coat, removed from a	nstructed using RNA isolated from untreated peripheral blood mononuclear cell d from buffy coat, removed from a 60-vear-old male.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	PE Biosystems, Foster City, CA.	
ABIPARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	PE Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	PE Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score GCG- specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

5

10

15

20

25

30

35

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

- a) an amino acid sequence selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43,
- b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:39, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43,
- c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43, and
- d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID

NO:41, SEQ ID NO:42, and SEQ ID NO:43.

- 2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43.
 - 3. An isolated polynucleotide encoding a polypeptide of claim 1.
 - 4. An isolated polynucleotide encoding a polypeptide of claim 2.

15

20

35

- 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86.
- 6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
 - 7. A cell transformed with a recombinant polynucleotide of claim 6.
- 30 8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
 - 9. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

5

10

15

20

25

30

- 10. An isolated antibody which specifically binds to a polypeptide of claim 1.
- 11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86,
- b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86.
 - c) a polynucleotide sequence complementary to a),
 - d) a polynucleotide sequence complementary to b), and
 - e) an RNA equivalent of a)-d).
- 12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

5

- 15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
 - 16. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

15

20

10

17. A pharmaceutical composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43.

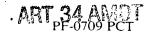
- 18. A method for treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.
- 19. A method for screening a compound for effectiveness as an agonist of a polypeptide of30 claim 1, the method comprising:
 - a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 - b) detecting agonist activity in the sample.
- 20. A pharmaceutical composition comprising an agonist compound identified by a method
 of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 20.

- 5 22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
 - a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 - b) detecting antagonist activity in the sample.

15

- 23. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.
 - 24. A method for treating a disease or condition associated with overexpression of functional TPPT, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 23.
 - 25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:
- a) combining the polypeptide of claim 1 with at least one test compound under suitable
 20 conditions, and
 - b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.
- 26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:
 - a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1.
 - b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.
 - 27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method



10

25

30

comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.
- 5 28. An isolated polynucleotide comprising a polynucleotide sequence of SEQ ID NO:83.
 - 29. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 28.
 - 30. A cell transformed with a recombinant polynucleotide of claim 29.
 - 31. A transgenic organism comprising a recombinant polynucleotide of claim 29.
- 32. A method for producing a polypeptide comprising an amino acid sequence of SEQ ID NO:40, the method comprising:
 - a) culturing the cell of claim 30 under conditions suitable for expression of the polypeptide, and
 - b) recovering the polypeptide so expressed.
- 33. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 28, the method comprising:
 - a) exposing a sample comprising the target polynucleotide to a compound, and
 - b) detecting altered expression of the target polynucleotide.
 - 34 An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
 - a) an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37,
 - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37,
 - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37, and
- of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37

5

- 35. An isolated polypeptide of claim 34 selected from the group consisting of SEQ ID NO 5. SEQ ID NO:11, and SEQ ID NO.37.
 - 36. An isolated polynucleotide encoding a polypeptide of claim 34.
 - 37. An isolated polynucleotide encoding a polypeptide of claim 35
- 38. An isolated polynucleotide of claim 37 selected from the group consisting of SEQ ID NO:48, SEQ ID NO:54, and SEQ ID NO:80.
 - 39. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 36.
- 40. A cell transformed with a recombinant polynucleotide of claim 39.
 - 41. A transgenic organism comprising a recombinant polynucleotide of claim 39.
 - 42. A method for producing a polypeptide of claim 34, the method comprising:
- a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 34, and
 - b) recovering the polypeptide so expressed.
 - 43. An isolated antibody which specifically binds to a polypeptide of claim 34.
 - 44. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
 - a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:48, SEQ ID NO:54, and SEQ ID NO:80,
 - b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48, SEQ ID NO:54, and SEQ ID NO:80.
 - c) a polynucleotide sequence complementary to a).
 - d) a polynucleotide sequence complementary to b), and
 - e) an RNA equivalent of a)-d).

25

10

15

20

25

- 45 An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 44.
- 46. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 44, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
 - 47. A method of claim 46, wherein the probe comprises at least 60 contiguous nucleotides.
- 48. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 44, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof
- 49. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 34 and a pharmaceutically acceptable excipient.
- 50 A pharmaceutical composition of claim 49, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37.
- 51. A method for treating a disease or condition associated with decreased expression of functional human transport proteins (TPPT), comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 49.
 - 52. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 347 the method comprising:
 - a) exposing a sample comprising a polypeptide of claim 34 to a compound, and
 - b) detecting agonist activity in the sample.

20

25

- 53. A pharmaceutical composition comprising an agonist compound identified by a method of claim 52 and a pharmaceutically acceptable excipient.
- 54 A method for treating a disease or condition associated with decreased expression of functional human transport proteins (TPPT), comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 53.
 - 55. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 34, the method comprising:
 - a) exposing a sample comprising a polypeptide of claim 34 to a compound, and
 - b) detecting antagonist activity in the sample.
- 56. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 55 and a pharmaceutically acceptable excipient.
 - 57. A method for treating a disease or condition associated with overexpression of functional human transport proteins (TPPT), comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 56.

58. A method of screening for a compound that specifically binds to the polypeptide of claim 34. said method comprising the steps of:

- a) combining the polypeptide of claim 34 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 34 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 34.
- 59. A method of screening for a compound that modulates the activity of the polypeptide of claim 34, said method comprising:
- a) combining the polypeptide of claim 34 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 34.
- b) assessing the activity of the polypeptide of claim 34 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 34 in the presence of the test compound with the activity of the polypeptide of claim 34 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 34 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 34.

	60. A method for screening a compound for effectiveness in altering expression of a target
	polynucleotide, wherein said target polynucleotide comprises a sequence of claim 38, the method
	comprising:
5	a) exposing a sample comprising the target polynucleotide to a compound, and
	b) detecting altered expression of the target polynucleotide.
	61. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:1.
10	62. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:2
	63. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO.3.
15	64. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:4.
73	65. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:6.
	66. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:7.
20	67. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:8.
	68. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:9.
25	69. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO.10.
- 5	70. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:12.
	71. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:13.
30	72. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:14.
	A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:15.
35	74. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:16.
ر ر	75. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:17.

	76. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:18
5	77. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:19.
J	78. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:20.
	79. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:21.
10	80. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:22.
	81. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:23.
15	82. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:24.
	83. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:25.
	84. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:26.
20	85. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:27.
	86 A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO.28.
25	87. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO.29.
	88 A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:30.
	89. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:31.
30	90. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:33.
	91. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:34.
- 35	92 A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:35.
	93. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:36.

	94. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:38
5	95 A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO.39
3	96. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:41
	97. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:43.
10	98 A method of claim 42, wherein the polypeptide has the sequence of SEQ ID NO:5.
	99. A method of claim 42, wherein the polypeptide has the sequence of SEQ ID NO:11.
15	100 A method of claim 42, wherein the polypeptide has the sequence of SEQ ID NO:37.
15	101. A diagnostic test for a condition or disease associated with the expression of human transport proteins (TPPT) in a biological sample comprising the steps of:
	a) combining the biological sample with an antibody of claim 10, under conditions suitable
	for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and
20	b) detecting the complex, wherein the presence of the complex correlates with the presence
	of the polypeptide in the biological sample.
	102. A diagnostic test for a condition or disease associated with the expression of human
	transport proteins (TPPT) in a biological sample comprising the steps of.
25	a) combining the biological sample with an antibody of claim 43, under conditions suitable
	for the antibody to bind the polypeptide and form an antibody:polypeptide complex: and
	b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
30	102. The weekle does of also as 100 to 100 t
30	103. The antibody of claim 10, wherein the antibody is:
	a) a chimeric antibody,
	b) a single chain antibody.
	c) a Fab fragment, or
- 35	d) a F(ab') ₂ fragment, or e) a humanized antibody.
22	~ Sy a numanized antibody,

10

15

20

- 104. The antibody of claim 43, wherein the antibody is:a) a chimeric antibody.b) a single chain antibody.c) a Fab fragment.
- d) a F(ab')₂ fragment, or
- e) a humanized antibody.
- 105. A composition comprising an antibody of claim 10 and an acceptable excipient.
- 106. A composition comprising an antibody of claim 43 and an acceptable excipient.
- 107. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 105.
- 108. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 106.
 - 109. A composition of claim 105, wherein the antibody is labeled
 - 110. A composition of claim 106, wherein the antibody is labeled
- 25 111. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 109.
 - 112. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 110
 - 113. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:
- the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6,

25

30

35

SEQ ID NO 7. SEQ ID NO:8. SEQ ID NO:9. SEQ ID NO:10. SEQ ID NO-12. SEQ ID NO.13. SEQ ID NO-14. SEQ ID NO:15. SEQ ID NO.16. SEQ ID NO:17. SEQ ID NO.18. SEQ ID NO:19. SEQ ID NO.20. SEQ ID NO:21. SEQ ID NO:22. SEQ ID NO:23. SEQ ID NO:24. SEQ ID NO:25. SEQ ID NO:26. SEQ ID NO:27. SEQ ID NO:28. SEQ ID NO:29. SEQ ID NO:30. SEQ ID NO:31. SEQ ID NO:33. SEQ ID NO:34. SEQ ID NO:35. SEQ ID NO:36. SEQ ID NO:38. SEQ ID NO:39. SEQ ID NO:41. and SEQ ID NO:43. or an immunogenic fragment thereof. under conditions to elicit an antibody response.

- b) isolating antibodies from said animal; and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:NO:1. SEQ ID NO.2. SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO 14. SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43.
- 20 114. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 43 comprising:
 - a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO 11, and SEQ ID NO:37, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - b) isolating antibodies from said animal, and
 - c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37.
 - 115. An antibody produced by a method of claim 113.
 - 116. An antibody produced by a method of claim 114.
 - 117. A composition comprising the antibody of claim 115 and a suitable carrier.
 - 118 A composition comprising the antibody of claim 116 and a suitable carrier.

10

15

20

- 119. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising.
- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2. SEQ ID NO:3. SEQ ID NO 4. SEQ ID NO:6. SEQ ID NO:7. SEQ ID NO:8. SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12. SEQ ID NO 13. SEQ ID NO:14. SEQ ID NO:15. SEQ ID NO:16. SEQ ID NO:17. SEQ ID NO:18. SEQ ID NO:19. SEQ ID NO:20. SEQ ID NO:21. SEQ ID NO:22. SEQ ID NO:23. SEQ ID NO:24, SEQ ID NO:25. SEQ ID NO:26. SEQ ID NO:27. SEQ ID NO:28. SEQ ID NO:29. SEQ ID NO:30. SEQ ID NO:31. SEQ ID NO:33. SEQ ID NO:34. SEQ ID NO:35. SEQ ID NO:36. SEQ ID NO:38. SEQ ID NO:39. SEQ ID NO:41. and SEQ ID NO:43. or an immunogenic fragment thereof, under conditions to elicit an antibody response;
 - b) isolating antibody producing cells from the animal:
 - c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells:
 - d) culturing the hybridoma cells; and
 - e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO 3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO.7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO.14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43.

25

30

- 120. A method of making a monoclonal antibody with the specificity of the antibody of claim 43 comprising:
- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5. SEQ ID NO 11, and SEQ ID NO:37, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
 - b) isolating antibody producing cells from the animal;
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;
 - d) culturing the hybridoma cells; and
- having an amino acid sequence selected from the group consisting of SEQ ID NO:5. SEQ ID NO:11, and SEQ ID NO:37.

15

20

25

30

121	A monoclonal	antibody	produced	by a	method	of claim	119
-----	--------------	----------	----------	------	--------	----------	-----

- 122. A monoclonal antibody produced by a method of claim 120
- 123. A composition comprising the antibody of claim 121 and a suitable carrier.
- 124. A composition comprising the antibody of claim 122 and a suitable carrier
- 10 125. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.
 - 126. The antibody of claim 43, wherein the antibody is produced by screening a Fab expression library.
 - 127. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.
 - 128. The antibody of claim 43, wherein the antibody is produced by screening a recombinant immunoglobulin library.
 - 129. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2. SEQ ID NO:3. SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO.7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO.13, SEQ ID NO:14. SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43 in a sample, comprising the steps of:
 - a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
 - b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO.7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22,

SEQ ID NO.23. SEQ ID NO.24. SEQ ID NO.25. SEQ ID NO.26. SEQ ID NO.27. SEQ ID NO 28. SEQ ID NO.29. SEQ ID NO:30. SEQ ID NO.31. SEQ ID NO:33. SEQ ID NO:34. SEQ ID NO:35. SEQ ID NO:36. SEQ ID NO:38. SEQ ID NO:39. SEQ ID NO:41. and SEQ ID NO 43 in the sample

5

- 130. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37 in a sample, comprising the steps of
- a) incubating the antibody of claim 43 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
 - b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37 in the sample.

15

20

25

30

10

- 131. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43 from a sample, the method comprising.
- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43.

35

132. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37 from a sample, the method comprising:

10

15

20

25

30

35

- a) incubating the antibody of claim 43 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO.5. SEQ ID NO:11, and SEQ ID NO.37.
- 133 A microarray wherein at least one element of the microarray is a polynucleotide of claim 12.

134. A microarray wherein at least one element of the microarray is a polynucleotide of claim 45.

- 135 A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:
 - a) labeling the polynucleotides of the sample.
 - b) contacting the elements of the microarray of claim 133 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
 - c) quantifying the expression of the polynucleotides in the sample.

136. A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:

- a) labeling the polynucleotides of the sample.
- b) contacting the elements of the microarray of claim 134 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
 - c) quantifying the expression of the polynucleotides in the sample
- 137. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, said target polynucleotide having a sequence of claim 11.
- 138. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, said target polynucleotide having a sequence of claim 44.

10

15

20

- 139 An array of claim 137, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide
- 140. An array of claim 138, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.
 - 141. An array of claim 139, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.
 - 142. An array of claim 140, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.
 - 143. An array of claim 137, which is a microarray.
 - 144. An array of claim 138, which is a microarray.
 - 145. An array of claim 139, further comprising said target polynucleotide hybridized to said first oligonucleotide or polynucleotide.
 - 146. An array of claim 140, further comprising said target polynucleotide hybridized to said first oligonucleotide or polynucleotide.
- 147. An array of claim 137, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.
 - 148. An array of claim 138, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.
- 30 149. An array of claim 137, wherein each distinct physical location on the substrate contains multiple nucleotide molecules having the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another physical location on the substrate.
- 35 150. An array of claim 138, wherein each distinct physical location on the substrate contains multiple nucleotide molecules having the same sequence, and each distinct physical location on the

substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another physical location on the substrate.

5 151. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1. 152 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.2 153 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3 10 154. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4. 155. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6. 15 156. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7. 157. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8. 158. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9. 20 150 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10. 160. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.12. 25 161 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13. 162 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14. 163. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15. 30 164 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16. 165 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17. - L66 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18. 35

	167 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO 19
5	168. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20
J	169. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO 21
	170. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.22.
10	171. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.23.
	172. A polypeptide of claim 1. comprising the amino acid sequence of SEQ ID NO.24.
15	173. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
	174 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.
	175. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.
20	176. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.
	177. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.
25	178 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.
	170. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.31.
	180. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.
30	181. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.
	182. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.
35	183. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.
	184. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

	185. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.39
5	186. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO 41
J	187. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.43
	188. A polypeptide of claim 34, comprising the amino acid sequence of SEQ ID NO.5
10	189. A polypeptide of claim 34, comprising the amino acid sequence of SEQ ID NO.11
	190. A polypeptide of claim 34, comprising the amino acid sequence of SEQ ID NO.37
15	191. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:44.
	192. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:45.
20	193. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:46.
	194. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO·47.
25	195. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:49.
30	196. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:50.
	197. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:51.
35	"198. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:52.

- 35

NO 65.

	NO.53		A polynucleotide of claim 11. comprising the polynucleotide sequence of	SEQ ID
5	NO.55		A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
10	NO:56		A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
10	NO:57		A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
15	NO:58		A polynucleotide of claim 11. comprising the polynucleotide sequence of	SEQ ID
	NO:59.		A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
20	NO:60	205. A	A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
25	NO.61.		A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
	NO:62	207 A	A polynucleotide of claim 11. comprising the polynucleotide sequence of	SEQ ID
30	NO:63	208. A	A polynucleotide of claim 11. comprising the polynucleotide sequence of	SEQ ID
	NO·64	209 A	A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID

· - 210 A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

- 35

NO:78.

	NO.66	211. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
5	NO:67.	212. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
10	NO:68	213. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
	NO:69.	214. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
15	NO:70	215. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
	NO:71.	216. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
20	NO:72	217. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
25	NO.73.	218 A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
23	NO:74	219. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
30	NO:76	220 A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID
	NO:77.	221. A polynucleotide of claim 11, comprising the polynucleotide sequence of	SEQ ID

- 222 A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

	223.	A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ II
NO:79.		

- 5 224. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:81.
 - 225. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:82.
- 226. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:84.
 - 227. A polynucleotide of claim 11. comprising the polynucleotide sequence of SEQ ID NO:86.
 - 228. A polynucleotide of claim 44, comprising the polynucleotide sequence of SEQ ID NO.48.
- 229. A polynucleotide of claim 44, comprising the polynucleotide sequence of SEQ IDNO:54.
 - 230. A polynucleotide of claim 44, comprising the polynucleotide sequence of SEQ ID NO:80.

(12) INTERNATIONAL AP

ATION PUBLISHED UNDER THE PATENT



(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 28 December 2000 (28.12.2000)

PCT

(10) International Publication Number WO 00/78953 A2

- (51) International Patent Classification7: C12N 15/12. 5/10, C07K 14/47, 14/705, 16/18, 16/28, C12Q 1/68, A61K 38/17, G01N 33/50, A01K 67/027
- (21) International Application Number: PCT/US00/16668
- (22) International Filing Date: 16 June 2000 (16.06.2000)
- (25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/139,923	17 June 1999 (17.06.1999)	US
60/148,177	10 August 1999 (10.08.1999)	US
60/149,357	18 August 1999 (18.08.1999)	US
60/162,287	28 October 1999 (28.10.1999)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/139,923 (CIP)
Filed on	17 June 1999 (17.06.1999)
US	60/148,177 (CIP)
Filed on	10 August 1999 (10.08.1999)
US	60/149,357 (CIP)
Filed on	18 August 1999 (18.08.1999)
US	60/162,287 (CIP)
Filed on	28 October 1999 (28.10.1999)

- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). YANG, Junming [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). HILLMAN, Jennifer, L.

[US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). BURFORD, Neil [GB/US]; 1308 4th Avenue, San Francisco, CA 94122 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rocksprings Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way, #1, Menlo Park, CA 94025 (US).

- (74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN TRANSPORT PROTEINS

(57) Abstract: The invention provides human transport proteins (TPPT) and polynucleotides which identify and encode TPPT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of TPPT.

Docket No.: PF-0709 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN TRANSPORT PROTEINS

/ X / is attached hereto. / / was filed on as application Serial No and if this box contains an X / /, was amended on / X / was filed as Patent Cooperation Treaty international application No. PCT/US00/16668 on 16 June, 2000, if this box contains an X /_ /, was amended on under Patent Cooperation Treaty Artical on 2001, and if this box contains an X /_ /, was amended on I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.	the specification of which:		
contains an X //, was amended on /X / was filed as Patent Cooperation Treaty international application No. PCT/US00/16668 o 16 June, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Articles on 2001, and if this box contains an X /_/, was amended on I hereby state that I have reviewed and understand the contents of the above-identified	/ X / is attached hereto.		
/X / was filed as Patent Cooperation Treaty international application No. PCT/US00/16668 of 16 June, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Artical 19 on 2001, and if this box contains an X /_/, was amended on I hereby state that I have reviewed and understand the contents of the above-identified	// was filed on	as application Serial No	and if this box
16 June, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Articles on 2001, and if this box contains an X /_/, was amended on I hereby state that I have reviewed and understand the contents of the above-identified	contains an X //, was amend	ed on	·
16 June, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Articles on 2001, and if this box contains an X /_/, was amended on I hereby state that I have reviewed and understand the contents of the above-identified	/X / was filed as Patent Coo	peration Treaty international application	No. PCT/US00/16668 on
I hereby state that I have reviewed and understand the contents of the above-identified		-	
•	19 on 2001, and if this	box contains an X /_/, was amended on	·
•	I hereby state that I have	e reviewed and understand the contents	of the above-identified
	<u>-</u>		

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Country	Number	Filing Date	Priority Claimed
			/_/ Yes /_/ No
			/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application		Status (Pending,
Serial No.	Filed	Abandoned, Patented)
60/139,923	June 17, 1999	Expired
60/148,177	August 10, 1999	Expired
60/149,357	August 18, 1999	Expired
60/162,287	October 28, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application		Status (Pending,
Serial No.	Filed	Abandoned, Patented)
		Troumbond Twombon,

I hereby appoint the following:

Lucy J. Billings	Reg. No. <u>36,749</u>
Michael C. Cerrone	Reg. No. <u>39,132</u>
Diana Hamlet-Cox	Reg. No. <u>33,302</u>
Richard C. Ekstrom	Reg. No. <u>37,027</u>
Barrie D. Greene	Reg. No. <u>46,740</u>
Lynn E. Murry	Reg. No. <u>42,918</u>
Shirley A. Recipon	Reg. No. <u>47,016</u>
Susan K. Sather	Reg. No. <u>44,316</u>
Michelle M. Stempien	Reg. No. <u>41,327</u>
David G. Streeter	Reg. No. <u>43,168</u>
Stephen Todd	Reg. No. <u>47,139</u>
P. Ben Wang	Reg. No. <u>41,420</u>

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

78718

LEGAL DEPARTMENT INCYTE GENOMICS, INC. 3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

First Joint Inventor: 1	Full name:	Preeti Lal
	Signature:	Preek al
	Date:	September 10, 2001
	Citizenship:	India
	Residence:	Santa Clara, California
	P.O. Address:	P.O. Box 5142 Santa Clara, California 95056
Second Joint Inventor: 2-	Full name:	Junming Yang
	Signature:	So of
	Date:	September 17,2001
	Citizenship:	China
	Residence:	San Jose, California
	P.O. Address:	7125 Bark Lane San Jose, California 95129

San Jose, California 95118

Third Joint Inventor: 3 - ∞	Full name:	Henry Yue
	Signature:	Janny Jue.
	Date:	September 24, 2001
	Citizenship:	United States
	Residence:	Sunnyvale, California
	P.O. Address:	826 Lois Avenue Sunnyvale, California 94087
Fourth Joint Inventor: 4	Full name:	Jennifer L. Hillman
	Signature:	Jul Hallun
	Date:	September 2 , 2001
	Citizenship:	United States
	Residence:	Mountain View, California
	P.O. Address:	230 Monroe Drive, #17 Mountain View, California 94040
Fifth Joint Inventor: 5 - co	Full name:	Y. Tom Tang
	Signature:	U.a.los
	Date:	Sept. 10 ,2001
	Citizenship:	United States
	Residence:	San Jose, California
	P.O. Address:	4230 Ranwick Court

Sixth Joint Inventor: 6 - co	Full name:	Olga Bandman
	Signature:	Olga Bacedneau 12 September, 2001
	Date:	12 september, 2001
	Citizenship:	United States
	Residence:	Mountain View, California
	P.O. Address:	366 Anna Avenue Mountain View, California 94043
Seventh Joint Inventor:	Full name:	Neil Burford
•	Signature:	
	Date:	
	Citizenship:	United States
	Residence:	Durham, Connecticut
	P.O. Address:	105 Wildwood Circle Durham, Connecticut 06422
Eighth Joint Inventor: 8 - cc	Full name:	Mariah R. Baughn
	Signature:	Mu P. Bh
	Date:	September 5, 2001
	Citizenship:	United States
	Residence:	San Leandro, California
	P.O. Address:	14244 Santiago Road San Leandro, California 94577

Ninth Joint Inventor: 9 - 9	Full name: Signature:	Yalda Azimzai All Azimzai All Azimzai
	Date: Citizenship:	United States 13, 2001
	Residence:	Castro Valley, California
	P.O. Address:	5518 Boulder Canyon Drive <u>Castro Valley, California 94552</u>
Tenth Joint Inventor: ↓ ○ - ∞	Full name:	Dyung Aina M. Lu
	Signature:	Jha le
	Date:	Opt 7,2001
	Citizenship:	United States
	Residence:	San Jose, California
	P.O. Address:	233 Coy Drive San Jose, California 95123
Eleventh Joint Inventor: 11 - C	· Full name:	Janice Au-Young
	Signature:	Janin ale peng
	Date:	September 7,2001
	Citizenship:	United States
	Residence:	Brisbane, California
	P.O. Address:	233 Golden Eagle Lane Brisbane, California 94005

MARTINE SERVICE CONTRACTOR OF THE SERVICE OF THE SE

Docket No.: PF-0709 USN

Twelfth Joint Inventor: 12- 00 Full name:

Chandra Patterson Arvizu 9/10/01

Signature:

Chandra awiga

Date:

September 10, 2001

Citizenship:

United States

Residence:

Menlo Park, California

P.O. Address:

490 Sherwood Way, #1

Menlo Park, California 94025

2.

Docket No.: PF-0709 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN TRANSPORT PROTEINS

the specification of which:
/ X / is attached hereto.
// was filed on as application Serial No and if this box contains an X //, was amended on
/X / was filed as Patent Cooperation Treaty international application No. PCT/US00/16668 on 16 June, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Articl 19 on 2001, and if this box contains an X /_/, was amended on
I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.
I acknowledge my duty to disclose information which is material to the examination of this

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the

1

application in accordance with Title 37, Code of Federal Regulations, \$1.56(a).

application for said subject matter the priority of which is claimed:

Country	Number	Filing Date	Priority Claimed
			/_/ Yes /_/ No
			/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application		Status (Pending,
Serial No.	Filed	Abandoned, Patented)
60/139,923	June 17, 1999	Expired
60/148,177	August 10, 1999	Expired
60/149,357	August 18, 1999	Expired
60/162,287	October 28, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application		Status (Pending,	
Serial No.	Filed	Abandoned, Patented)	
I hereby appo	oint the following:		
Lucy J. Billings		Reg. No. 36,749	
Michael C. Cerrone		Reg. No. 39,132	
Diana Hamlet-Cox		Reg. No. 33,302	
Richard C. Ekstrom		Reg. No. 37,027	
Barrie D. Greene		Reg. No. 46,740	
Lynn E. Murry		Reg. No. 42,918	
Shirley A. Recipon		Reg. No. 47,016	
Susan K. Sather		Reg. No. 44,316	
Michelle M. Stempie	en	Reg. No. 41,327	
David G. Streeter		Reg. No. 43,168	
Stephen Todd		Reg. No. 47,139	
P. Ben Wang		Reg. No. 41,420	

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

LEGAL DEPARTMENT INCYTE GENOMICS, INC. 3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

First Joint Inventor:	Full name:	Preeti Lal
	Signature:	
	Date:	, 2001
	Citizenship:	India
	Residence:	Santa Clara, California
	P.O. Address:	P.O. Box 5142 Santa Clara, California 95056
Second Joint Inventor:	Full name:	Junming Yang
	Signature:	
	Date:	, 2001
	Citizenship:	China
	Residence:	San Jose, California
	P.O. Address:	7125 Bark Lane San Jose, California 95129

Third Joint Inventor:	Full name:	Henry Yue
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	Sunnyvale, California
	P.O. Address:	826 Lois Avenue Sunnyvale, California 94087
Fourth Joint Inventor:	Full name:	Jennifer L. Hillman
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	Mountain View, California
	P.O. Address:	230 Monroe Drive, #17 Mountain View, California 94040
Fifth Joint Inventor:	Full name:	Y. Tom Tang
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	San Jose, California
	P.O. Address:	4230 Ranwick Court San Jose, California 95118

Sixth Joint Inventor:	Full name:	Olga Bandman
	Signature:	
	Date:	
	Citizenship:	United States
	Residence:	Mountain View, California
	P.O. Address:	366 Anna Avenue Mountain View, California 94043
Seventh Joint Inventor: ←-∞	Full name:	Neil Burford
	Signature:	New Mill.
	Date:	9/21, 2001
	Citizenship:	9/21/01 NB. United States UNITED WINGSOM
	Residence:	Durham, Connecticut
	P.O. Address:	105 Wildwood Circle Durham, Connecticut 06422
Eighth Joint Inventor:	Full name:	Mariah R. Baughn
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	San Leandro, California
	P.O. Address:	14244 Santiago Road San Leandro, California 94577

Ninth Joint Inventor:	Full name:	Yalda Azimzai
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	Castro Valley, California
	P.O. Address:	5518 Boulder Canyon Drive <u>Castro Valley, California</u> 94552
Tenth Joint Inventor:	Full name:	Dyung Aina M. Lu
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	San Jose, California
	P.O. Address:	233 Coy Drive San Jose, California 95123
Eleventh Joint Inventor:	Full name:	Janice Au-Young
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	Brisbane, California
	P.O. Address:	233 Golden Eagle Lane Brisbane, California 94005

Fwelfth Joint Inventor:	Full name:	Chandra Patterson
	Signature:	
	Date:	, 2001
	Citizenship:	United States
	Residence:	Menlo Park, California
	P.O. Address:	490 Sherwood Way, #1 Menlo Park, California, 94025

WO 00/78953 PCT/US00/16668

SEQUENCE LISTING

```
<110> INCYTE GENOMICS, INC.
      LAL, Preeti
      YANG, Junming
      YUE, Henry
      HILLMAN, Jennifer L.
      TANG, Y. Tom
      BANDMAN, Olga
BURFORD, Neil
      BAUGHN, Mariah R.
      AZIMZAI, Yalda
      LU, Dyung Aina M.
      AU-YOUNG, Janice
      PATTERSON, Chandra
<120> HUMAN TRANSPORT PROTEINS
<130> PF-0709 PCT
<140> To Be Assigned
<141> Herewith
<150> 60/139,923; 60/148,177; 60/149,357; 60/162,287
<151> 1999-06-17; 1999-08-10; 1999-08-18; 1999-10-28
<160> 86
<170> PERL Program
<210> 1
<211> 623
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 264114CD1
<400> 1
Met Ser Thr Gln Asp Glu Arg Gln Ile Asn Thr Glu Tyr Ala Val
                                       10
                                                            15
Ser Leu Leu Glu Gln Leu Lys Leu Phe Tyr Glu Gln Gln Leu Phe
                 20
                                                            30
Thr Asp Ile Val Leu Ile Val Glu Gly Thr Glu Phe Pro Cys His
                 35
                                       40
Lys Met Val Leu Ala Thr Cys Ser Ser Tyr Phe Arg Ala Met Phe
                 50
                                       55
                                                            60
Met Ser Gly Leu Ser Glu Ser Lys Gln Thr His Val His Leu Arg
                 65
                                       70
                                                            75
Asn Val Asp Ala Ala Thr Leu Gln Ile Ile Ile Thr Tyr Ala Tyr
                 80
                                       85
                                                            90
Thr Gly Asn Leu Ala Met Asn Asp Ser Thr Val Glu Gln Leu Tyr
                 95
                                      100
                                                           105
Glu Thr Ala Cys Phe Leu Gln Val Glu Asp Val Leu Gln Arg Cys
                110
                                      115
                                                           120
Arg Glu Tyr Leu Ile Lys Lys Ile Asn Ala Glu Asn Cys Val Arg
                125
                                      130
                                                           135
Leu Leu Ser Phe Ala Asp Leu Phe Ser Cys Glu Glu Leu Lys Gln
                140
                                      145
                                                           150
Ser Ala Lys Arg Met Val Glu His Lys Phe Thr Ala Val Tyr His
                155
                                      160
Gln Asp Ala Phe Met Gln Leu Ser His Asp Leu Leu Ile Asp Ile
                170
                                      175
                                                           180
Leu Ser Ser Asp Asn Leu Asn Val Glu Lys Glu Glu Thr Val Arg
                185
                                      190
                                                           195
Glu Ala Ala Met Leu Trp Leu Glu Tyr Asn Thr Glu Ser Arg Ser
```

WO 00/78953 PCT/US00/16668

```
200
                                      205
Gln Tyr Leu Ser Ser Val Leu Ser Gln Ile Arg Ile Asp Ala Leu
                 215
                                      220
                                                          225
Ser Glu Val Thr Gln Arg Ala Trp Phe Gln Gly Leu Pro Pro Asn
                 230
                                      235
                                                          240
Asp Lys Ser Val Val Val Gln Gly Leu Tyr Lys Ser Met Pro Lys
                 245
                                      250
                                                          255
Phe Phe Lys Pro Arg Leu Gly Met Thr Lys Glu Glu Met Met
                                                          Ile
                 260
                                      265
                                                          270
Phe Ile Glu Ala Ser Ser Glu Asn Pro Cys Ser Leu Tyr Ser Ser
                 275
                                      280
                                                          285
Val Cys Tyr Ser Pro Gln Ala Glu Lys Val Tyr Lys Leu Cys Ser
                 290
                                      295
                                                          300
Pro Pro Ala Asp Leu His Lys Val Gly Thr Val Val Thr Pro Asp
                 305
                                      310
                                                          315
Asn Asp Ile Tyr Ile Ala Gly Gly Gln Val Pro Leu Lys Asn Thr
                 320
                                      325
                                                          330
Lys Thr Asn His Ser Lys Thr Ser Lys Leu Gln Thr Ala Phe Arg
                 335
                                      340
                                                          345
Thr Val Asn Cys Phe Tyr Trp Phe Asp Ala Gln Gln Asn Thr Trp
                 350
                                      355
                                                          360
Phe Pro Lys Thr Pro Met Leu Phe Val Arg Ile Lys Pro Ser Leu
                 365
                                      370
                                                          375
Val Cys Cys Glu Gly
                     Tyr Ile Tyr Ala Ile Gly Gly Asp Ser
                                                          Val
                 380
                                      385
                                                          390
Gly Gly Glu Leu Asn Arg Arg Thr Val Glu Arg Tyr Asp Thr Glu
                 395
                                      400
                                                          405
Lys Asp Glu Trp Thr Met Val Ser Pro Leu Pro Cys Ala Trp Gln
                 410
                                      415
                                                          420
Trp Ser Ala Ala Val Val His Asp Cys Ile Tyr Val Met Thr
                 425
                                      430
                                                           435
Leu Asn Leu Met Tyr Cys Tyr Phe Pro Arg Ser Asp Ser Trp Val
                 440
                                      445
                                                          450
Glu Met Ala Met Arg Gln Thr Ser Arg Ser Phe Ala Ser Ala Ala
                 455
                                      460
                                                          465
Ala Phe Gly Asp Lys Ile Phe Tyr Ile Gly Gly Leu His Ile Ala
                 470
                                      475
                                                          480
Thr Asn Ser Gly Ile Arg Leu Pro Ser Gly Thr Val Asp Gly Ser
                 485
                                      490
                                                          495
Ser Val Thr Val Glu Ile Tyr Asp Val Asn Lys Asn Glu Trp Lys
                 500
                                      505
                                                          510
Met Ala Ala Asn Ile Pro Ala Lys Arg Tyr Ser Asp Pro Cys Val
                 515
                                      520
                                                          525
Arg Ala Val Val Ile Ser Asn Ser Leu Cys Val Phe Met Arg Glu
                 530
                                      535
                                                          540
Thr His Leu Asn Glu Arg Ala Lys Tyr Val Thr Tyr Gln Tyr Asp
                 545
                                      550
                                                          555
Leu Glu Leu Asp Arg Trp Ser Leu Arg Gln His Ile Ser Glu Arg
                 560
                                      565
                                                          570
Val Leu Trp Asp Leu Gly Arg Asp Phe Arg Cys Thr Val Gly Lys
                 575
                                      580
                                                          585
Leu Tyr Pro Ser Cys Leu Glu Glu Ser Pro Trp Lys Pro Pro Thr
                 590
                                      595
                                                          600
Tyr Leu Phe Ser Thr Asp Gly Thr Glu Glu Phe Glu Leu Asp
                                                          Gly
                 605
                                      610
                                                          615
Glu Met Val Ala Leu Pro Pro Val
                 620
<210> 2
<211> 99
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 1455669CD1
```

Met Ala Ala Pro Ala Glu Pro Cys Ala Gly Gln Gly Val Trp Asn

WO 00/78953 PCT/US00/16668

```
10
                                                           15
Gln Thr Glu Pro Glu Pro Ala Ala Thr Ser Leu Leu Ser Leu Cys
                                      25
                 20
Phe Leu Arg Thr Ala Gly Val Trp Val Pro Pro Met Tyr Leu Trp
                 35
                                      40
                                                           45
Val Leu Gly Pro Ile Tyr Leu Leu Phe Ile His His Gly Arg
                 50
                                      55
                                                           60
Gly Tyr Leu Arg Met Ser Pro Leu Phe Lys Ala Lys Met Val Ala
                 65
                                                           75
Ala Ile Pro Gly Ser Leu Glu Pro Gly Asn Val Arg Gly Arg Gln
                 80
Gly Thr Gly Trp Asn Leu Val Lys Ser
                 95
<210> 3
<211> 374
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2084989CD1
<400> 3
Met Glu Ser Lys Met Gly Glu Leu Pro Leu Asp Ile Asn Ile Gln
                  5
                                      10
Glu Pro Arg Trp Asp Gln Ser Thr Phe Leu Gly Arg Ala Arg His
                 20
                                      25
                                                           30
Phe Phe Thr Val Thr Asp Pro Arg Asn Leu Leu Leu Ser Gly Ala
                 35
                                      40
                                                           45
Gln Leu Glu Ala Ser Arg Asn Ile Val Gln Asn Tyr Arg Ala Gly
                 50
                                      55
                                                           60
Val Val Thr Pro Gly Ile Thr Glu Asp Gln Leu Trp Arg Ala Lys
                 65
                                      70
Tyr Val Tyr Asp Ser Ala Phe His Pro Asp Thr Gly Glu Lys Val
                 80
                                      85
                                                           90
Val Leu Ile Gly Arg Met Ser Ala Gln Val Pro Met Asn Met Thr
                 95
                                     100
                                                          105
Ile Thr Gly Cys Met Leu Thr Phe Tyr Arg Lys Thr Pro Thr Val
                110
                                     115
                                                          120
Val Phe Trp Gln Trp Val Asn Gln Ser Phe Asn Ala Ile Val Asn
                125
                                     130
                                                          135
Tyr Ser Asn Arg Ser Gly Asp Thr Pro Ile Thr Val Arg Gln Leu
                140
                                     145
                                                          150
Gly Thr Ala Tyr Val Ser Ala Thr Thr Gly Ala Val Ala Thr Ala
                155
                                     160
                                                          165
Leu Gly Leu Lys Ser Leu Thr Lys His Leu Pro Pro Leu Val Gly
                170
                                     175
                                                          180
Arg Phe Val Pro Phe Ala Ala Val Ala Ala Ala Asn Cys Ile Asn
                185
                                     190
                                                          195
Ile Pro Leu Met Arg Gln Arg Glu Leu Gln Val Gly Ile Pro Val
                200
                                     205
                                                          210
Ala Asp Glu Ala Gly Gln Arg Leu Gly Tyr Ser Val Thr Ala Ala
                215
                                     220
                                                          225
Lys Gln Gly Ile Phe Gln Val Val Ile Ser Arg Ile Cys Met Ala
                230
                                     235
                                                          240
Ile Pro Ala Met Ala Ile Pro Pro Leu Ile Met Asp Thr Leu Glu
                245
                                     250
                                                          255
Lys Lys Asp Phe Leu Lys Val Gly Asp
                                     Cys Thr Ser Leu Val Leu
                260
                                     265
                                                          270
Glu Trp Ala Met Ala Gly Arg Ser Asp Gln Ala Pro Thr Leu Ser
                275
                                     280
                                                          285
Pro Ala Ser Pro Asp Ser Leu Arg Leu Ala Ser Pro Ser Pro Asp
                290
                                     295
                                                          300
Pro Cys Thr Ala Ser Ser Thr Phe Val His Ser Ala Arg Met Asn
                305
                                     310
Trp Ala Gly Val Lys Glu Leu Cys Arg Gly Arg Arg Gly Gln
                320
                                     325
                                                          330
Arg Lys Glu Thr Asn Phe Ile Ser Val Thr Pro Val Ala Ser Asp
```

```
335
                                     340
Thr Gln Lys Gly Thr Val Ile Val Met Leu Asp Leu Met Leu Ile
                350
                                     355
                                                         360
Leu Leu Pro Pro Ser Ala Ser Ile Leu Arg Gly Thr His Gly
                365
<210> 4
<211> 271
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2501034CD1
<400> 4
Met Gly Asn Gly Gly Arg Ser Gly Leu Gln Gln Gly Lys Gly Asn
                                                          15
Val Asp Gly Val Ala Ala Thr Pro Thr Ala Ala Ser Ala Ser Cys
                 20
                                       25
                                                           30
Gln Tyr Arg Cys Ile Glu Cys Asn Gln Glu Ala Lys Glu Leu Tyr
                 35
                                      40
                                                           45
Arg Asp Tyr Asn His Gly Val Leu Lys Ile Thr Ile Cys Lys Ser
                 50
                                      55
                                                           60
Cys Gln Lys Pro Val Asp Lys Tyr Ile Glu Tyr Asp Pro Val Ile
                  65
                                       70
                                                           75
Ile Leu Ile Asn Ala Ile Leu Cys Lys Ala Gln Ala Tyr Arg His
                 80
                                      85
Ile Leu Phe Asn Thr Gln Ile Asn Ile His Gly Lys Leu Cys Ile
                 95
                                     100
                                                          105
Phe Cys Leu Leu Cys Glu Ala Tyr Leu Arg Trp Trp Gln Leu Gln
                110
                                     115
                                                          120
Asp Ser Asn Gln Asn Thr Ala Pro Asp Asp Leu Ile Arg Tyr Ala
                125
                                     130
                                                          135
Lys Glu Trp Asp Phe Tyr Arg Met Phe Ala Ile Ala Ala Leu Glu
                140
                                     145
                                                          150
Gln Thr Ala Tyr Phe Ile Gly Ile Phe Thr Phe Leu Trp Val Glu
                155
                                      160
                                                          165
Arg Pro Met Thr Ala Lys Lys Lys Pro Asn Phe Ile Leu Leu Leu
                170
                                      175
                                                          180
Lys Ala Leu Leu Ser Ser Tyr Gly Lys Leu Leu Ile Pro
                185
                                      190
                                                          195
Ala Val Ile Trp Glu His Asp Tyr Thr Ser Val Cys Leu Lys Leu
                200
                                      205
                                                          210
Ile Lys Val Phe Val Leu Thr Ser Asn Phe Gln Ala Ile Arg Val
                 215
                                      220
                                                          225
Thr Leu Asn Ile Asn Arg Lys Leu Ser Phe Leu Ala Val Leu Ser
                230
                                      235
Gly Leu Leu Glu Ser Ile Met Val Tyr Phe Phe Gln Ser Met
                245
                                      250
                                                          255
Glu Trp Asp Val Gly Ser Asp Tyr Ala Ile Phe Lys Ser Gln Asp
                260
                                      265
Phe
<210> 5
<211> 323
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2745212CD1
<400> 5
Met Ala Pro Lys Gln Asp Pro Lys Pro Lys Phe Gln Glu Gly Glu
                                       10
                                                           15
Arg Val Leu Cys Phe His Gly Pro Leu Leu Tyr Glu Ala Lys Cys
                  20
                                       25
Val Lys Val Ala Ile Lys Asp Lys Gln Val Lys Tyr Phe Ile His
```

```
Tyr Ser Gly Trp Asn Lys Asn Trp Asp Glu Trp Val Pro Glu Ser
                 50
                                      55
Arg Val Leu Lys Tyr Val Asp Thr Asn Leu Gln Lys Gln Arg Glu
                 65
                                      70
Leu Gln Lys Ala Asn Gln Glu Gln Tyr Ala Glu Gly Lys Met Arg
                 80
                                      85
                                                           90
Gly Ala Ala Pro Gly Lys Lys Thr Ser Gly Leu Gln Gln Lys Asn
                 95
                                     100
                                                          105
Val Glu Val Lys Thr Lys Lys Asn Lys Gln Lys Thr Pro Gly Asn
                110
                                     115
                                                          120
Gly Asp Gly Gly Ser Thr Ser Glu Thr Pro Gln Pro Pro Arg Lys
                125
                                     130
                                                          135
Lys Arg Ala Arg Val Asp Pro Thr Val Glu Asn Glu Glu Thr Phe
                140
                                     145
                                                          150
Met Asn Arg Val Glu Val Lys Val Lys Ile Pro Glu Glu Leu Lys
                155
                                     160
                                                          165
Pro Trp Leu Val Asp Asp Trp Asp Leu Ile Thr Arg Gln Lys Gln
                170
                                     175
                                                          180
Leu Phe Tyr Leu Pro Ala Lys Lys Asn Val Asp Ser Ile Leu Glu
                185
                                     190
                                                          195
Asp Tyr Ala Asn Tyr Lys Lys Ser Arg Gly Asn Thr Asp Asn Lys
                200
                                     205
                                                          210
Glu Tyr Ala Val Asn Glu Val Val Ala Gly Ile Lys Glu Tyr Phe
                215
                                     220
Asn Val Met Leu Gly Thr Gln Leu Leu Tyr Lys Phe Glu Arg Pro
                230
                                     235
                                                          240
Gln Tyr Ala Glu Ile Leu Ala Asp His Pro Asp Ala Pro Met Ser
                245
                                     250
                                                          255
Gln Val Tyr Gly Ala Pro His Leu Leu Arg Leu Phe Val Arg Ile
                260
                                     265
                                                          270
Gly Ala Met Leu Ala Tyr Thr Pro Leu Asp Glu Lys Ser Leu Ala
                275
                                     280
Leu Leu Leu Asn Tyr Leu His Asp Phe Leu Lys Tyr Leu Ala Lys
                290
                                     295
Asn Ser Ala Thr Leu Phe Ser Ala Ser Asp Tyr Glu Val Ala Pro
                305
                                     310
                                                          315
Pro Glu Tyr His Arg Lys Ala Val
                320
<210> 6
<211> 274
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 4833111CD1
<400> 6
Met Asp Arg Pro Gly Phe Val Ala Ala Leu Val Ala Gly Gly Val
  1
                                      10
                                                           15
Ala Gly Val Ser Val Asp Leu Ile Leu Phe Pro Leu Asp Thr Ile
                 20
                                      25
Lys Thr Arg Leu Gln Ser Pro Gln Gly Phe Ser Lys Ala Gly Gly
                 35
                                      40
Phe His Gly Ile Tyr Ala Gly Val Pro Ser Ala Ala Ile Gly Ser
                 50
                                      55
                                                           60
Phe Pro Asn Ala Ala Ala Phe Phe Ile Thr Tyr Glu Tyr Val Lys
                 65
                                      70
                                                           75
Trp Phe Leu His Ala Asp Ser Ser Ser Tyr Leu Thr Pro Met Lys
                 80
                                      85
                                                           90
His Met Leu Ala Ala Ser Ala Gly Glu Val Val Ala Cys Leu Ile
                 95
                                     100
                                                          105
Arg Val Pro Ser Glu Val Val Lys Gln Arg Ala Gln Val Ser Ala
                110
                                     115
                                                          120
Ser Thr Arg Thr Phe Gln Ile Phe Ser Asn Ile Leu Tyr Glu Glu
                125
                                    130
Gly Ile Gln Gly Leu Tyr Arg Gly Tyr Lys Ser Thr Val Leu Arg
```

```
140
                                      145
Glu Ile Pro Phe Ser Leu Val Gln Phe Pro Leu Trp Glu Ser Leu
                 155
                                      160
                                                          165
Lys Ala Leu Trp Ser Trp Arg Gln Asp His Val Val Asp Ser Trp
                 170
                                      175
                                                          180
Gln Ser Ala Val Cys Gly Ala Phe Ala Gly Gly Phe Ala Ala
                 185
                                      190
                                                          195
Val Thr Thr Pro Leu Asp Val Ala Lys Thr Arg Ile Thr Leu Ala
                 200
                                      205
                                                          210
Lys Ala Gly Ser Ser Thr Ala Asp Gly Asn Val Leu Ser Val Leu
                 215
                                      220
                                                          225
His Gly Val Trp Arg Ser Gln Gly Leu Ala Gly Leu Phe Ala Gly
                 230
                                      235
                                                          240
Val Phe Pro Arg Met Ala Ala Ile Ser Leu Gly Gly Phe Ile Phe
                 245
                                      250
                                                          255
Leu Gly Ala Tyr Asp Arg Thr His Ser Leu Leu Leu Glu Val Gly
                 260
                                      265
                                                          270
Arg Lys Ser Pro
<210> 7
<211> 291
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 876677CD1
<400> 7
Met Asp Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys Glu Asn
  1
                                       10
Phe Val Gly Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp Ile
                  20
                                       25
Leu Phe Ser Leu Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile
                  35
                                       40
                                                           45
Ser Ile Trp Met Cys Leu Lys Ile Ile Lys Glu Tyr Glu Arg Ala
                  50
                                       55
Val Val Phe Arg Leu Gly Arg Ile Gln Ala Asp Lys Ala Lys Gly
                  65
                                       70
                                                           75
Pro Gly Leu Ile Leu Val Leu Pro Cys Ile Asp Val Phe Val Lys
                  ឧ೧
                                       85
Val Asp Leu Arg Thr Val Thr Cys Asn Ile Pro Pro Gln Glu Ile
                  95
                                      100
                                                          105
Leu Thr Arg Asp Ser Val Thr Thr Gln Val Asp Gly Val Val Tyr
                 110
                                      115
                                                          120
Tyr Arg Ile Tyr Ser Ala Val Ser Ala Val Ala Asn Val Asn Asp
                 125
                                      130
                                                          135
Val His Gln Ala Thr Phe Leu Leu Ala Gln Thr Thr Leu Arg Asn
                 140
                                      145
                                                          150
Val Leu Gly Thr Gln Thr Leu Ser Gln Ile Leu Ala Gly Arg Glu
                 155
                                      160
                                                          165
Glu Ile Ala His Ser Ile Gln Thr Leu Leu Asp Asp Ala Thr Glu
                 170
                                      175
                                                          180
Leu Trp Gly Ile Arg Val Ala Arg Val Glu Ile Lys Asp Val Arg
                 185
                                      190
                                                          195
Ile Pro Val Gln Leu Gln Arg Ser Met Ala Ala Glu Ala Glu Ala
                 200
                                      205
                                                          210
Thr Arg Glu Ala Arg Ala Lys Val Leu Ala Ala Glu Gly Glu Met
                 215
                                      220
                                                          225
Asn Ala Ser Lys Ser Leu Lys Ser Ala Ser Met Val Leu Ala Glu
                 230
                                      235
                                                          240
Ser Pro Ile Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu Ser Thr
                 245
                                      250
                                                          255
Val Ala Thr Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro Met
                 260
                                      265
                                                          270
Asn Ile Leu Glu Gly Ile Gly Gly Val Ser Tyr Asp Asn His Lys
                 275
                                      280
                                                          285
Lys Leu Pro Asn Lys Ala
```

<210> 8

<211> 381 <212> PRT <213> Homo sapiens <220> <221> misc_feature <223> Incyte ID No: 2326143CD1 <400> 8 Met Ser Arg His Glu Gly Val Ser Cys Asp Ala Cys Leu Lys Gly Asn Phe Arg Gly Arg Arg Tyr Lys Cys Leu Ile Cys Tyr Asp Leu Cys Ala Ser Cys Tyr Glu Ser Gly Ala Thr Thr Thr Arg His Thr Thr Asp His Pro Met Gln Cys Ile Leu Thr Arg Val Asp Phe Asp Leu Tyr Tyr Gly Gly Glu Ala Phe Ser Val Glu Gln Pro Gln Ser Phe Thr Cys Pro Tyr Cys Gly Lys Met Gly Tyr Thr Glu Thr Ser Leu Gln Glu His Val Thr Ser Glu His Ala Glu Thr Ser Thr Glu Val Ile Cys Pro Ile Cys Ala Ala Leu Pro Gly Gly Asp Pro Asn His Val Thr Asp Asp Phe Ala Ala His Leu Thr Leu Glu His Arg Ala Pro Arg Asp Leu Asp Glu Ser Ser Gly Val Arg His Val Arg Arg Met Phe His Pro Gly Arg Gly Leu Gly Gly Pro Arg Ala Arg Arg Ser Asn Met His Phe Thr Ser Ser Ser Thr Gly Gly Leu Ser Ser Ser Gln Ser Ser Tyr Ser Pro Ser Asn Arg Glu Ala Met Asp Pro Ile Ala Glu Leu Leu Ser Gln Leu Ser Gly Val Arg Arg Ser Ala Gly Gly Gln Leu Asn Ser Ser Gly Pro Ser Ala Ser Gln Leu Gln Gln Leu Gln Met Gln Leu Gln Leu Glu Arg Gln His Ala Gln Ala Ala Arg Gln Gln Leu Glu Thr Ala Arg Asn Ala Thr Arg Arg Thr Asn Thr Ser Ser Val Thr Thr Thr Ile Thr Gln Ser Thr Ala Thr Thr Asn Ile Ala Asn Thr Glu Ser Ser Gln Gln Thr Leu Gln Asn Ser Gln Phe Leu Leu Thr Arg Leu Asn Asp Pro Lys Met Ser Glu Thr Glu Arg Gln Ser Met Glu Ser Glu Arg Ala Asp Arg Ser Leu Phe Val Gln Glu Leu Leu Leu Ser Thr Leu Val Arg Glu Glu Ser Ser Ser Asp Glu Asp Asp Arg Gly Glu Met Ala Asp Phe Gly Ala Met Gly Cys Val Asp Ile Met Pro Leu Asp Val Ala Leu Glu Asn Leu Asn Leu Lys Glu Ser Asn Lys Gly Asn Glu Pro Pro Pro Pro Leu <210> 9 <211> 190 <212> PRT <213> Homo sapiens

<220>

```
<221> misc_feature
<223> Incyte ID No: 2786302CD1
Met Lys Tyr Gly Asn Glu Ile Met Asn Lys Asp Pro Val Phe Arg
  1
                                      10
                                                           15
Ile Ser Pro Arg Ser Arg Glu Thr His Pro Asn Pro Glu Glu Pro
                  20
                                       25
                                                           30
Glu Glu Glu Asp Glu Asp Val Gln Ala Glu Arg Val Gln Ala Ala
                  35
                                       40
                                                           45
Asn Ala Leu Thr Ala Pro Asn Leu Glu Glu Glu Pro Val Ile Thr
                  50
                                       55
Ala Ser Cys Leu His Lys Glu Tyr Tyr Glu Thr Lys Lys Ser Cys
                  65
                                       70
                                                           75
Phe Ser Thr Arg Lys Lys Lys Ile Ala Ile Arg Asn Val Ser Phe
                  80
                                       85
                                                           90
Cys Val Lys Lys Gly Glu Val Leu Gly Leu Leu Gly His Asn Gly
                  95
                                      100
                                                          105
Ala Gly Lys Ser Thr Ser Ile Lys Met Ile Thr Gly Cys Thr Lys
                 110
                                     115
                                                          120
Pro Thr Ala Gly Val Val Leu Gln Gly Ser Arg Ala Ser Val
                 125
                                      130
                                                          135
Arg Gln Gln His Asp Asn Ser Leu Lys Phe Leu Gly Tyr Cys Pro
                 140
                                      145
                                                          150
Gln Glu Asn Ser Leu Trp Pro Lys Leu Thr Met Lys Glu His Leu
                 155
                                      160
                                                          165
Glu Leu Tyr Ala Ala Val Glu Arg Leu Gly Gln Lys Arg Cys Cys
                 170
                                      175
                                                          180
Ser Gln Tyr Phe Thr Ile Gly Gly Arg
                                     Ser
                 185
<210> 10
<211> 297
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 3735780CD1
<400> 10
Met Met Asp Ser Glu Ala His Glu Lys Arg Pro Pro Ile Leu Thr
                                       10
Ser Ser Lys Gln Asp Ile Ser Pro His
                                     Ile
                                         Thr Asn Val Gly
                                                         Glu
                  20
                                       25
                                                           30
Met Lys His Tyr Leu Cys Gly Cys Cys Ala Ala Phe Asn Asn Val
                  35
                                       40
                                                           45
Ala Ile Thr Phe Pro Ile Gln Lys Val Leu Phe Arg Gln Gln Leu
                  50
                                       55
                                                           60
Tyr Gly Ile Lys Thr Arg Asp Ala Ile Leu Gln Leu Arg Arg Asp
                  65
                                       70
Gly Phe Arg Asn Leu Tyr Arg Gly Ile Leu Pro Pro Leu Met Gln
                  80
                                       85
                                                           90
Lys Thr Thr Leu Ala Leu Met Phe Gly
                                         Leu Tyr Glu Asp Leu
                  95
                                      100
                                                          105
Ser Cys Leu Leu His Lys His Val Ser Ala Pro Glu Phe Ala Thr
                 110
                                      115
                                                          120
Ser Gly Val Ala Ala Val Leu Ala Gly Thr Thr Glu Ala Ile Phe
                 125
                                      130
                                                          135
Thr Pro Leu Glu Arg Val Gln Thr Leu Leu Gln Asp His Lys His
                 140
                                      145
                                                          150
His Asp Lys Phe Thr Asn Thr Tyr Gln Ala Phe Lys Ala Leu Lys
                 155
                                      160
                                                          165
Cys His Gly Ile Gly Glu Tyr Tyr Arg Gly Leu Val Pro Ile Leu
                 170
                                      175
                                                          180
Phe Arg Asn Gly Leu Ser Asn Val Leu Phe Phe Gly Leu Arg Gly
                 185
                                      190
                                                          195
Pro Ile Lys Glu His Leu Pro Thr Ala Thr Thr His Ser Ala His
```

```
200
                                     205
                                                          210
Leu Val Asn Asp Phe Ile Cys Gly Gly Leu Leu Gly Ala Met Leu
                215
                                     220
Gly Phe Leu Phe Phe Pro Ile Asn Val Val Lys Thr Arg Ile Gln
                230
                                     235
                                                          240
Ser Gln Ile Gly Gly Glu Phe Gln Ser Phe Pro Lys Val Phe Gln
                245
                                     250
                                                          255
Lys Ile Trp Leu Glu Arg Asp Arg Lys Leu Ile Asn Leu Phe Arg
                260
                                     265
                                                          270
Gly Ala His Leu Asn Tyr His Arg Ser Leu Ile Ser Trp Gly Ile
                275
                                     280
Ile Asn Ala Thr Tyr Glu Phe Leu Leu Lys Val Ile
                290
                                     295
<210> 11
<211> 89
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 039026CD1
<400> 11
Met Ala Ala Gln Ile Pro Glu Ser Asp Gln Ile Lys Gln Phe Lys
  1
                                      10
                                                           15
Glu Phe Leu Gly Thr Tyr Asn Lys Leu Thr Glu Thr Cys Phe Leu
                 20
                                      25
                                                           30
Asp Cys Val Lys Asp Phe Thr Thr Arg Glu Val Lys Pro Glu Glu
                 35
                                      40
                                                           45
Thr Thr Cys Ser Glu His Cys Leu Gln Lys Tyr Leu Lys Met Thr
                 50
                                      55
Gln Arg Ile Ser Met Arg Phe Gln Glu Tyr His Ile Gln Gln Asn
                 65
                                      70
Glu Ala Leu Ala Ala Lys Ala Gly Leu Leu Gly Gln Pro Arg
                 80
                                      85
<210> 12
<211> 115
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 260607CD1
<400> 12
Met Ala Leu Ile Pro Ser Arg Val Trp Leu Pro Phe Ala Val Trp
                                      10
                                                           15
Val Val Asp Ser Ala Pro Val Arg Gly Leu Val Arg Arg Glu Pro
                 20
                                      25
                                                           30
Phe Leu Arg Thr Gly Ser Phe Ile Ala Leu Phe Tyr Phe Pro Pro
                 35
Leu Leu Pro Val Leu Ile Asn Leu Phe Ser Phe Phe Leu Thr Pro
                 50
                                      55
                                                           60
Ser Phe Trp Arg Gln Leu Gly Ala Ile Leu Val Tyr Ala Ser Leu
                 65
                                      70
                                                           75
Leu Ala Glu Lys Thr Pro Phe Lys Thr Gln Arg Thr Leu Glu Gly
                 80
                                      85
                                                           90
Asp Ala Leu Val Gly Ser Val Ser Ile Phe Leu Cys Ala Lys Asp
                 95
                                     100
                                                          105
Arg Gln Thr Glu Ala Glu Arg Gly Cys Ser
                 110
<210> 13
<211> 675
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
```

<223> Incyte ID No: 1429651CD1

<400> 13 Met Glu Ser Gly Thr Ser Ser Pro Gln Pro Pro Gln Leu Asp Pro Leu Asp Ala Phe Pro Gln Lys Gly Leu Glu Pro Gly Asp Ile Ala Val Leu Val Leu Tyr Phe Leu Phe Val Leu Ala Val Gly Leu Trp Ser Thr Val Lys Thr Lys Arg Asp Thr Val Lys Gly Tyr Phe Leu Ala Gly Gly Asp Met Val Trp Trp Pro Val Gly Ala Ser Leu Phe Ala Ser Asn Val Gly Ser Gly His Phe Ile Gly Leu Ala Gly Ser Gly Ala Ala Thr Gly Ile Ser Val Ser Ala Tyr Glu Leu Asn Gly Leu Phe Ser Val Leu Met Leu Ala Trp Ile Phe Leu Pro Ile Tyr Ile Ala Gly Gln Val Thr Thr Met Pro Glu Tyr Leu Arg Lys Arg Phe Gly Gly Ile Arg Ile Pro Ile Ile Leu Ala Val Leu Tyr Leu Phe Ile Tyr Ile Phe Thr Lys Ile Ser Val Asp Met Tyr Ala Gly Ala Ile Phe Ile Gln Gln Ser Leu His Leu Asp Leu Tyr Leu Ala Ile Val Gly Leu Leu Ala Ile Thr Ala Val Tyr Thr Val Ala Gly Gly Leu Ala Ala Val Ile Tyr Thr Asp Ala Leu Gln Thr Leu Ile Met Leu Ile Gly Ala Leu Thr Leu Met Gly Tyr Ser Phe Ala Ala Val Gly Gly Met Glu Gly Leu Lys Glu Lys Tyr Phe Leu Ala Leu Ala Ser Asn Arg Ser Glu Asn Ser Ser Cys Gly Leu Pro Arg Glu Asp Ala Phe His Ile Phe Arg Asp Pro Leu Thr Ser Asp Leu Pro Trp Pro Gly Val Leu Phe Gly Met Ser Ile Pro Ser Leu Trp Tyr Trp Cys Thr Asp Gln Val Ile Val Gln Arg Thr Leu Ala Ala Lys Asn Leu Ser His Ala Lys Gly Gly Ala Leu Met Ala Ala Tyr Leu Lys Val Leu Pro Leu Phe Ile Met Val Phe Pro Gly Met Val Ser Arg Ile Leu Phe Pro Asp Gln Val Ala Cys Ala Asp Pro Glu Ile Cys Gln Lys Ile Cys Ser Asn Pro Ser Gly Cys Ser Asp Ile Ala Tyr Pro Lys Leu Val Leu Glu Leu Leu Pro Thr Gly Leu Arg Gly Leu Met Met Ala Val Met Val Ala Ala Leu Met Ser Ser Leu Thr Ser Ile Phe Asn Ser Ala Ser Thr Ile Phe Thr Met Asp Leu Trp Asn His Leu Arg Pro Arg Ala Ser Glu Lys Glu Leu Met Ile Val Gly Arg Val Phe Val Leu Leu Val Leu Val Ser Ile Leu Trp Ile Pro Val Val Gln Ala Ser Gln Gly Gly Gln Leu Phe Ile Tyr Ile Gln Ser Ile Ser Ser Tyr Leu Gln Pro Pro Val Ala Val Val Phe Ile Met Gly Cys Phe Trp Lys Arg Thr Asn Glu Lys Gly Ala Phe Trp Gly Leu Ile Ser Gly Leu Leu Leu Gly Leu Val Arg Leu

```
485
                                     490
                                                          495
Val Leu Asp Phe Ile Tyr Val Gln Pro Arg Cys Asp Gln Pro Asp
                 500
                                     505
                                                          510
Glu Arg Pro Val Leu Val Lys Ser Ile His Tyr Leu Tyr Phe Ser
                 515
                                     520
                                                          525
Met Ile Leu Ser Thr Val Thr Leu Ile Thr Val Ser Thr Val Ser
                 530
                                     535
Trp Phe Thr Glu Pro Pro Ser Lys Glu Met Val Ser His Leu Thr
                 545
                                     550
                                                          555
Trp Phe Thr Arg His Asp Pro Val Val Gln Lys Glu Gln Ala Pro
                 560
                                     565
                                                          570
Pro Ala Ala Pro Leu Ser Leu Thr Leu Ser Gln Asn Gly Met Pro
                 575
                                     580
                                                          585
Glu Ala Ser Ser Ser Ser Val Gln Phe Glu Met Val Gln Glu
                 590
                                     595
                                                          600
Asn Thr Ser Lys Thr His Ser Cys Asp Met Thr Pro Lys Gln Ser
                 605
                                     610
                                                          615
Lys Val Val Lys Ala Ile Leu Trp Leu Cys Gly Ile Gln Glu Lys
                 620
                                     625
Gly Lys Glu Glu Leu Pro Ala Arg Ala Glu Ala Ile Ile Val Ser
                 635
                                     640
                                                          645
Leu Glu Glu Asn Pro Leu Val Lys Thr Leu Leu Asp Val Asn Leu
                 650
                                     655
                                                          660
Ile Phe Cys Val Ser Cys Ala Ile Phe Ile Trp Gly Tyr Phe Ala
                 665
                                     670
<210> 14
<211> 320
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2069971CD1
<400> 14
Met Tyr His Cys His Ser Gly Ser Lys Pro Thr Glu Lys Gly Ala
  1
                                      10
Asn Glu Tyr Ala Tyr
                    Ala Lys Trp Lys Leu Cys Ser Ala Ser Ala
                  20
                                      25
                                                           30
Ile Cys Phe Ile Phe Met Ile Ala Glu Val Val Gly Gly His Ile
                  35
                                       40
Ala Gly Ser Leu Ala Val Val Thr Asp Ala Ala His Leu Leu Ile
                  50
                                      55
                                                           60
Asp Leu Thr Ser Phe Leu Leu Ser Leu Phe Ser Leu Trp Leu Ser
                  65
                                      70
Ser Lys Pro Pro Ser Lys Arg Leu Thr Phe Gly Trp His Arg Ala
                  80
                                      85
Glu Ile Leu Gly Ala Leu Leu Ser Ile Leu Cys Ile Trp Val Val
                  95
                                     100
                                                          105
Thr Gly Val Leu Val Tyr Leu Ala Cys Glu Arg Leu Leu Tyr Pro
                 110
                                     115
                                                          120
Asp Tyr Gln Ile Gln Ala Thr Val Met Ile Ile Val Ser Ser Cys
                 125
                                     130
                                                          135
Ala Val Ala Ala Asn Ile Val Leu Thr Val Val Leu His Gln Arg
                 140
                                     145
                                                          150
Cys Leu Gly His Asn His Lys Glu Val Gln Ala Asn Ala Ser Val
                 155
                                     160
                                                          165
Arg Ala Ala Phe Val His Ala Leu Gly Asp Leu Phe Gln Ser Ile
                 170
                                     175
                                                          180
Ser Val Leu Ile Ser Ala Leu Ile Ile Tyr Phe Lys Pro Glu Tyr
                 185
                                     190
                                                          195
Lys Ile Ala Asp Pro Ile Cys Thr Phe Ile Phe Ser Ile Leu Val
                 200
                                     205
                                                          210
Leu Ala Ser Thr Ile Thr Ile Leu Lys Asp Phe Ser Ile Leu Leu
                 215
                                     220
Met Glu Gly Val Pro Lys Ser Leu Asn Tyr Ser Gly Val Lys Glu
                 230
                                     235
                                                          240
```

```
Leu Ile Leu Ala Val Asp Gly Val Leu Ser Val His Ser Leu His
                245
                                     250
                                                          255
Ile Trp Ser Leu Thr Met Asn Gln Val Ile Leu Ser Ala His Val
                260
                                     265
                                                          270
Ala Thr Ala Ala Ser Arg Asp Ser Gln Val Val Arg Arg Glu Ile
                275
                                     280
                                                          285
Ala Lys Ala Leu Ser Lys Ser Phe Thr Met His Ser Leu Thr
                                                          Tle
                290
                                     295
                                                          300
Gln Met Glu Ser Pro Val Asp Gln Asp Pro Asp Cys Leu Phe Cys
                305
                                     310
                                                          315
Glu Asp Pro Cys Asp
                320
<210> 15
<211> 462
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2329339CD1
<400> 15
Met Ala Glu Glu Gln Glu Phe Thr Gln Leu Cys Lys Leu Pro Ala
  1
                                      10
                                                           15
Gln Pro Ser His Pro His Cys Val Asn Asn Thr Tyr Arg Ser Ala
                 20
                                      25
                                                            30
Gln His Ser Gln Ala Leu Leu Arg Gly Leu Leu Ala Leu Arg Asp
                 35
                                      40
                                                            45
Ser Gly Ile Leu Phe Asp Val Val Leu Val Val Glu Gly Arg His
                 50
                                      55
                                                            60
Ile Glu Ala His Arg Ile Leu Leu Ala Ala Ser Cys Asp Tyr Phe
                 65
                                      70
                                                            75
Arg Gly Met Phe Ala Gly Gly Leu Lys Glu Met Glu Gln Glu Glu
                 80
                                      85
                                                           90
Val Leu Ile His Gly Val Ser Tyr Asn Ala Met Cys Gln Ile Leu
                 95
                                     100
                                                          105
His Phe Ile Tyr Thr Ser Glu Leu Glu Leu Ser Leu Ser Asn Val
                110
                                     115
                                                          120
Gln Glu Thr Leu Val Ala Ala Cys Gln Leu Gln Ile Pro Glu Ile
                125
                                     130
                                                          135
Ile His Phe Cys Cys Asp Phe Leu Met Ser Trp Val Asp Glu Glu
                140
                                     145
                                                          150
Asn Ile Leu Asp Val
                    Tyr Arg Leu Ala Glu Leu Phe Asp Leu Ser
                155
                                     160
                                                          165
Arg Leu Thr Glu Gln Leu Asp Thr Tyr Ile Leu Lys Asn Phe Val
                170
                                     175
                                                          180
Ala Phe Ser Arg Thr Asp Lys Tyr Arg Gln Leu Pro Leu Glu Lys
                185
                                     190
                                                          195
Val Tyr Ser Leu Leu Ser Ser Asn Arg Leu Glu Val Ser Cys Glu
                200
                                     205
                                                          210
Thr Glu Val Tyr Glu Gly Ala Leu Leu Tyr His Tyr Ser Leu Glu
                215
                                     220
                                                          225
Gln Val Gln Ala Asp Gln Ile Ser Leu His Glu Pro Pro Lys Leu
                230
                                     235
                                                          240
Leu Glu Thr Val Arg
                    Phe Pro Leu Met Glu Ala Glu Val Leu Gln
                245
                                     250
                                                          255
Arg Leu His Asp Lys Leu Asp Pro Ser Pro Leu Arg Asp Thr Val
                260
                                     265
                                                          270
Ala Ser Gly Leu Met Tyr His Arg Asn Glu Ser Leu Gln Pro Ser
                275
                                     280
                                                          285
Leu Gln Ser Pro Gln Thr Glu Leu Arg Ser Asp Phe Gln Cys Val
                290
                                     295
                                                          300
Val Gly Phe Gly Gly Ile His Ser Thr Pro Ser Thr Val Leu Ser
                305
                                     310
                                                          315
Asp Gln Ala Lys Tyr Leu Asn Pro Leu Leu Gly Glu Trp Lys His
                320
                                     325
                                                          330
Phe Thr Ala Ser Leu Ala Pro Arg Met Ser Asn Gln Gly Ile Ala
                335
                                     340
                                                          345
```

```
Val Leu Asn Asn Phe Val Tyr Leu Ile Gly Gly Asp Asn Asn Val
                350
                                     355
Gln Gly Phe Arg Ala Glu Ser Arg Cys Trp Arg Tyr Asp Pro Arg
                365
                                     370
                                                          375
His Asn Arg Trp Phe Gln Ile Gln Ser Leu Gln Gln Glu His Ala
                380
                                     385
                                                          390
Asp Leu Ser Val Cys Val Val Gly Arg Tyr Ile Tyr Ala Val Ala
                395
                                     400
                                                          405
Gly Arg Asp Tyr His Asn Asp Leu Asn Ala Val Glu Arg Tyr Asp
                410
                                     415
                                                           420
Pro Ala Thr Asn Ser Trp Ala Tyr Val Ala Pro Leu Lys Arg Glu
                425
                                     430
                                                          435
Val Tyr Ala His Ala Gly Ala Thr Leu Glu Gly Lys Met Tyr Ile
                440
                                     445
                                                          450
Thr Cys Gly Arg Lys Leu Ile Pro Phe Ser Glu Gly
                455
                                     460
<210> 16
<211> 98
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2540219CD1
<400> 16
Met Arg Ala Cys Ala Val Trp Leu Ala Gly Gly Met Ala Gly Ala
                                      10
Ile Ser Trp Gly Thr Ala Thr Pro Met Asp Val Val Lys Ser Arg
                 20
                                      25
                                                           30
Leu Gln Ala Asp Gly Val Tyr Leu Asn Lys Tyr Lys Gly Val Leu
                 35
                                      40
                                                            45
Asp Cys Ile Ser Gln Ser Tyr Gln Lys Glu Gly Leu Lys Val Phe
                 50
                                      55
                                                           60
Phe Arg Gly Ile Thr Val Asn Ala Val Arg Gly Phe Pro Met Ser
                 65
Ala Ala Met Phe Leu Gly Tyr Glu Leu Ser Leu Gln Ala Ile Arg
                 80
                                      85
Gly Asp His Ala Val Thr Ser Pro
                 95
<210> 17
<211> 748
<212> PRT
<213> Homo sapiens
<220>
<221> misć_feature
<223> Incyte ID No: 2722462CD1
Met Asn Tyr Gln Glu Ala Ala Ile Tyr Leu Gln Glu Gly Glu Asn
                                      10
                                                           15
Asn Asp Lys Phe Phe Thr His Pro Lys Asp Ala Lys Ala Leu Ala
                 20
                                      25
                                                           30
Ala Tyr Leu Phe Ala His Asn His Leu Phe Tyr Leu Met Glu Leu
                 35
                                      40
                                                            45
Ala Thr Ala Leu Leu Leu Leu Leu Ser Leu Cys Glu Ala Pro
                 50
                                      55
                                                            60
Ala Val Pro Ala Leu Arg Leu Gly Ile Tyr Val His Ala Thr Leu
                 65
                                       70
                                                            75
Glu Leu Phe Ala Leu Met Val Val Val Phe Glu Leu Cys Met Lys
                 80
                                      85
                                                            90
Leu Arg Trp Leu Gly Leu His Thr Phe Ile Arg His Lys Arg Thr
                 95
                                      100
                                                           105
Met Val Lys Thr Ser Val Leu Val Val Gln Phe Val Glu Ala Ile
                110
                                     115
                                                          120
Val Val Leu Val Arg Gln Met Ser His Val Arg Val Thr Arg Ala
                125
                                     130
                                                          135
```

_	_			_										
Leu	Arg	Cys	Ile	Phe 140	Leu	Val	Asp	Cys	Arg 145	Tyr	Cys	Gly	Gly	Val 150
Arg	Arg	Asn	Leu	Arg 155	Gln	Ile	Phe	Gln	Ser 160	Leu	Pro	Pro	Phe	Met 165
Asp	Ile	Leu	Leu		Leu	Leu	Phe	Phe		Ile	Ile	Phe	Ala	
Leu	Gly	Phe	Tyr	Leu 185	Phe	Ser	Pro	Asn	Pro 190	Ser	Asp	Pro	Tyr	Phe 195
Ser	Thr	Leu	Glu	Asn 200	Ser	Ile	Val	Ser		Phe	Val	Leu	Leu	Thr 210
Thr	Ala	Asn	Phe		Asp	Val	Met	Met		Ser	Tyr	Ser	Arg	Asn 225
Pro	Trp	Ser	Cys		Phe	Phe	Ile	Val		Leu	Ser	Ile	Glu	Leu 240
Tyr	Phe	Ile	Met		Leu	Leu	Leu	Ala	Val	Val	Phe	Asp	Thr	Phe
Asn	Asp	Ile	Glu		Arg	Lys	Phe	Lys		Leu	Leu	Leu	His	
Arg	Thr	Ala	Ile		His	Ala	Tyr	Arg		Leu	Ile	Ser	Gln	
Arg	Pro	Ala	Gly		Ser	Tyr	Arg	Gln		Glu	Gly	Leu	Met	
Phe	Tyr	Lys	Pro		Met	Ser	Ala	Arg		Arg	Tyr	Leu	Thr	
Lys	Ala	Leu	Asn		Asn	Asn	Thr	Pro		Leu	Ser	Leu	Lys	
Phe	Tyr	Asp	Ile		Glu	Val	Ala	Ala		Lys	Trp	Lys	Ala	
Lys	Asn	Arg	Glu		Trp	Phe	Asp	Glu		Pro	Arg	Thr	Ala	
Leu	Ile	Phe	Lys		Ile	Asn	Ile	Leu		Lys	Ser	Lys	Ala	
Gln	Tyr	Phe	Met		Leu	Val	Val	Ala		Asn	Gly	Val	Trp	
Leu	Val	Glu	Thr		Met	Leu	Lys	Gly		Asn	Phe	Phe	Ser	
His	Val	Pro	Trp		Tyr	Leu	Val	Phe		Thr	Ile	Tyr	Gly	
Glu	Leu	Phe	Leu		Val	Ala	Gly	Leu	415 Gly 430	Pro	Val	Glu	Tyr	
Ser	Ser	Gly	Trp		Leu	Phe	Asp	Phe		Val	Thr	Val	Phe	
Phe	Leu	Gly	Leu		Ala	Leu	Ala	Leu	Asn	Met	Glu	Pro	Phe	
Phe	Ile	Val	Val		Arg	Pro	Leu	Gln	460 Leu 475	Leu	Arg	Leu	Phe	
Leu	Lys	Glu	Arg		Arg	Asn	Val	Leu	Asp	Thr	Met	Phe	Glu	
Leu	Pro	Arg	Met		Ser	Leu	Gly	Leu	490 Thr	Leu	Leu	Ile	Phe	
Tyr	Ser	Phe	Ala	Ile	Val	Gly	Met	Glu	505 Phe	Phe	Cys	Gly	Ile	
Phe	Pro	Asn	Cys		Asn	Thr	Ser	Thr		Ala	Asp	Ala	Tyr	
Trp	Arg	Asn	His	530 Thr	Val	Gly	Asn	Arg		Val	Val	Glu	Glu	
Tyr	Tyr	Tyr	Leu	545 Asn	Asn	Phe	Asp	Asn		Leu	Asn	Ser	Phe	
Thr	Leu	Phe	Glu	560 Leu	Thr	Val	Val	Asn		Trp	Tyr	Ile	Ile	
Glu	Gly	Val	Thr	575 Ser	Gln	Thr	Ser	His		Ser	Arg	Leu	Tyr	585 Phe
Met	Thr	Phe	Tyr	590 Ile	Val	Thr	Met	Val		Met	Thr	Ile	Ile	600 Val
Ala	Phe	Ile	Leu	605 Glu	Ala	Phe	Val	Phe		Met	Asn	Tyr	Ser	615 Arg
				620 Ser	Glu				625					630
				635					640					645

```
Glu Ile Ser Lys Glu Glu Leu Val Ala Val Leu Glu Leu Tyr Arg
                650
                                    655
Glu Ala Arg Gly Ala Ser Ser Asp Val Thr Arg Leu Leu Glu Thr
                665
                                    670
                                                         675
Leu Ser Gln Met Glu Arg Tyr Gln Gln His Ser Met Val Phe Leu
                680
                                    685
                                                         690
Gly Arg Arg Ser Arg Thr Lys Ser Asp Leu Ser Leu Lys Met
                                                         Tyr
                695
                                    700
                                                         705
Gln Glu Glu Ile Gln Glu Trp Tyr Glu Glu His Ala Arg Glu Gln
                710
                                    715
                                                         720
Glu Gln Gln Arg Gln Leu Ser Ser Ala Ala Pro Ala Ala Gln
                725
                                    730
                                                         735
Gln Pro Pro Gly Ser Arg Gln Arg Ser Gln Thr Val Thr
                                    745
<210> 18
<211> 507
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2739264CD1
<400> 18
Met Ala Phe Asn Phe Gly Ala Pro Ser Gly Thr Ser Gly Thr Ala
                                     10
Ala Ala Thr Ala Ala Pro Ala Gly Gly Phe Gly Phe Gly Thr
                 20
                                     25
                                                          30
Thr Ser Thr Thr Ala Gly Ser Ala Phe Ser Phe Ser Ala Pro Thr
                 35
                                     40
Asn Thr Gly Thr Thr Gly Leu Phe Gly Gly Thr Gln Asn Lys Gly
                 50
                                     55
                                                          60
Phe Gly Phe Gly Thr Gly Phe Gly Thr Thr Gly Thr Ser
                                                         Thr
                 65
                                     70
                                                          75
Gly Leu Gly Thr Gly Leu Gly Thr Gly Leu Gly Phe Gly Gly
                                                         Phe
                 80
Asn Thr Gln Gln Gln Gln Thr Thr Leu Gly Gly Leu Phe Ser
                 95
                                    100
                                                         105
Gln Pro Thr Gln Ala Pro Thr Gln Ser Asn Gln Leu Ile Asn Thr
               110
                                    115
                                                         120
Ala Ser Ala Leu Ser Ala Pro Thr Leu Leu Gly Asp Glu Arg Asp
                125
                                    130
                                                         135
Ala Ile Leu Ala Lys Trp Asn Gln Leu Gln Ala Phe Trp Gly
                                                         Thr
                140
                                    145
                                                          150
Gly Lys Gly Tyr Phe Asn Asn Ile Pro Pro Val Glu Phe Thr
                155
                                    160
                                                          165
Gln Glu Asn Pro Phe Cys Arg Phe Lys Ala Val Gly Tyr Ser
                                                         Cvs
                170
                                    175
                                                          180
Met Pro Ser Asn Lys Asp Glu Asp Gly Leu Val Val Leu Val Phe
                185
                                    190
                                                          195
Asn Lys Lys Glu Thr Glu Ile Arg Ser Gln Gln Gln Leu Val
                200
                                    205
                                                         210
Glu Ser Leu His Lys Val Leu Gly Gly Asn Gln Thr Leu Thr
                                                         Val
                215
                                    220
                                                          225
Asn Val Glu Gly Thr Lys Thr Leu Pro Asp Asp Gln Thr Glu Val
                230
                                    235
                                                          240
Val Ile Tyr Val Val Glu Arg Ser Pro Asn Gly Thr Ser Arg Arg
                245
                                    250
Val Pro Ala Thr Thr Leu Tyr Ala His Phe Glu Gln Ala Asn Ile
                260
                                    265
                                                          270
Lys Thr Gln Leu Gln Gln Leu Gly Val Thr Leu Ser Met Thr Arg
                275
                                    280
                                                          285
Thr Glu Leu Ser Pro Ala Gln Ile Lys Gln Leu Leu Gln Asn Pro
                290
                                    295
                                                          300
Pro Ala Gly Val Asp Pro Ile Ile Trp Glu Gln Ala Lys Val Asp
                305
                                    310
                                                          315
Asn Pro Asp Ser Glu Lys Leu Ile Pro Val Pro Met Val Gly
                                                         Phe
```

```
Lys Glu Leu Leu Arg Arg Leu Lys Val Gln Asp Gln Met Thr Lys
                 335
                                      340
                                                          345
Gln His Gln Thr Arg Leu Asp Ile Ile Ser Glu Asp Ile Ser Glu
                 350
                                      355
                                                          360
Leu Gln Lys Asn Gln Thr Thr Ser Val Ala Lys Ile Ala Gln Tyr
                 365
                                      370
                                                          375
Lys Arg Lys Leu Met Asp Leu Ser His Arg Thr Leu Gln Val Leu
                 380
                                      385
                                                          390
Ile Lys Gln Glu Ile Gln Arg Lys Ser Gly Tyr Ala Ile Gln Ala
                 395
                                      400
                                                          405
Asp Glu Glu Gln Leu Arg Val Gln Leu Asp Thr Ile Gln Gly Glu
                 410
                                      415
                                                          420
Leu Asn Ala Pro Thr Gln Phe Lys Gly Arg Leu Asn Glu Leu Met
                 425
                                      430
                                                          435
Ser Gln Ile Arg Met Gln Asn His Phe Gly Ala Val Arg Ser Glu
                 440
                                      445
                                                          450
Glu Arg Tyr Tyr Ile Asp Ala Asp Leu Leu Arg Glu Ile Lys Gln
                 455
                                      460
                                                          465
His Leu Lys Gln Gln Gln Glu Gly Leu Ser His Leu Ile Ser Ile
                 470
                                      475
                                                          480
Ile Lys Asp Asp Leu Glu Asp Ile Lys Leu Val Glu His Gly Leu
                 485
                                      490
                                                          495
Asn Glu Thr Ile His Ile Arg Gly Gly Val Phe Ser
                 500
<210> 19
<211> 592
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2758310CD1
<400> 19
Met Trp Phe Cys Gly Gln Ser Thr Pro Phe Gly Cys Glu Leu His
  1
Asp Thr Cys Val Gln Leu Cys His Phe His Ser Ala Leu Leu His
                 20
                                       25
                                                           30
Arg Arg Gln Lys Pro Trp Pro Ser Pro Ala Val Phe Phe Arg Arg
                  35
                                       40
                                                           45
Asn Val Arg Gly Leu Pro Pro Arg Phe Ser Ser Pro Thr Pro Leu
                                       55
                                                           60
Trp Arg Lys Val Leu Ser Thr Ala Val Val Gly Ala Pro Leu Leu
                  65
                                       70
                                                           75
Leu Gly Ala Arg Tyr
                    Val Met Ala Glu Ala Arg Glu Lys Arg Arg
                  80
                                       85
                                                           90
Met Arg Leu Val Val Asp Gly Met Gly Arg Phe Gly Arg Ser Leu
                  95
                                      100
                                                          105
Lys Val Gly Leu Gln Ile Ser Leu Asp Tyr Trp Trp Cys Thr Asn
                110
                                      115
                                                          120
Val Val Leu Arg Gly Trp Lys Ser Pro Gly Tyr Leu Glu Val Met
                 125
                                      130
                                                          135
Ser Ala Cys His Gln Arg Ala Ala Asp Ala Leu Val Ala Gly Ala
                 140
                                                          150
                                      145
Ile Ser Asn Gly Gly Leu Tyr Val Lys Leu Gly Gln Gly Leu Cys
                 155
                                      160
                                                          165
Ser Phe Asn His Leu Leu Pro Pro Glu Tyr Thr Arg Thr Leu Arg
                170
                                      175
                                                          180
Val Leu Glu Asp Arg Ala Leu Lys Arg Gly Phe Gln Glu Val Asp
                185
                                      190
                                                          195
Glu Leu Phe Leu Glu Asp Phe Gln Ala Leu Pro His Glu Leu Phe
                 200
                                      205
                                                          210
Gln Glu Phe Asp Tyr Gln Pro Ile Ala Ala Ala Ser Leu Ala Gln
                 215
                                      220
                                                          225
Val His Arg Ala Lys Leu His Asp Gly Thr Ser Val Ala Val Lys
                 230
                                      235
                                                          240
Val Gln Tyr Ile Asp Leu Arg Asp Arg Phe Asp Gly Asp Ile His
                 245
                                      250
                                                          255
```

```
Thr Leu Glu Leu Leu Arg Leu Val Glu Val Met His Pro Ser
                260
                                    265
Phe Gly Phe Ser Trp Val Leu Gln Asp Leu Lys Gly Thr Leu Ala
                275
                                    280
                                                          285
Gln Glu Leu Asp Phe Glu Asn Glu Gly Arg Asn Ala Glu Arg Cys
                290
                                    295
                                                          300
Ala Arg Glu Leu Ala His Phe Pro Tyr Val Val Val Pro Arg
                305
                                    310
                                                          315
His Trp Asp Lys Ser Ser Lys Arg Val Leu Thr Ala Asp Phe Cys
                320
                                    325
                                                          330
Ala Gly Cys Lys Val Asn Asp Val Glu Ala Ile Arg Ser Gln Gly
                335
                                    340
                                                          345
Leu Ala Val His Asp Ile Ala Glu Lys Leu Ile Lys Ala Phe Ala
                350
                                    355
                                                          360
Glu Gln Ile Phe Tyr Thr Gly Phe Ile His Ser Asp Pro His Pro
                365
                                    370
                                                          375
Gly Asn Val Leu Val Arg Lys Gly Pro Asp Gly Lys Ala Glu Leu
                380
                                    385
                                                          390
Val Leu Leu Asp His Gly Leu Tyr Gln Phe Leu Glu Glu Lys Asp
                395
                                    400
                                                          405
Arg Ala Ala Leu Cys Gln Leu Trp Arg Ala Ile Ile Leu Arg Asp
                410
                                    415
                                                          420
Asp Ala Ala Met Arg Ala His Ala Ala Leu Gly Val Gln Asp
                425
                                    430
                                                          435
Tyr Leu Leu Phe Ala Glu Met Leu Met Gln Arg Pro Val Arg Leu
                440
                                    445
                                                          450
Gly Gln Leu Trp Gly Ser His Leu Leu Ser Arg Glu Glu Ala Ala
                455
                                    460
                                                          465
Tyr Met Val Asp Met Ala Arg Glu Arg Phe Glu Ala Val Met Ala
                470
                                     475
Val Leu Arg Glu Leu Pro Arg Pro Met Leu Leu Val Leu Arg Asn
                485
                                    490
                                                          495
Ile Asn Thr Val Arg Ala Ile Asn Val Ala Leu Gly Ala Pro Val
                500
                                    505
                                                          510
Asp Arg Tyr Phe Leu Met Ala Lys Arg Ala Val Arg Gly Trp Ser
                515
                                    520
                                                          525
Arg Leu Ala Gly Ala Thr Tyr Arg Gly Val Tyr Gly Thr Ser Leu
                530
                                     535
                                                          540
Leu Arg His Ala Lys Val Val Trp Glu Met Leu Lys Phe Glu Val
                545
                                     550
                                                          555
Ala Leu Arg Leu Glu Thr Leu Ala Met Arg Leu Thr Ala Leu Leu
                560
                                     565
                                                          570
Ala Arg Ala Leu Val His Leu Ser Leu Val Pro Pro Ala Glu Glu
                575
                                     580
                                                          585
Leu Tyr Gln Tyr Leu Glu Thr
                590
<210> 20
<211> 841
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2762348CD1
<400> 20
Met Ala Ser Val Phe Arg Ser Glu Glu Met Cys Leu Ser Gln Leu
                                                           15
                                     10
Phe Leu Gln Val Glu Ala Ala Tyr Cys Cys Val Ala Glu Leu Gly
                 20
                                                           30
                                      25
Glu Leu Gly Leu Val Gln Phe Lys Asp Leu Asn Met Asn Val Asn
                 35
                                      40
                                                           45
Ser Phe Gln Arg Lys Phe Val Asn Glu Val Arg Arg Cys Glu Ser
                 50
                                      55
Leu Glu Arg Ile Leu Arg Phe Leu Glu Asp Glu Met Gln Asn Glu
                 65
                                      70
                                                           75
Ile Val Val Gln Leu Leu Glu Lys Ser Pro Leu Thr Pro Leu Pro
                 80
                                      85
```

Arg	Glu	Met	Ile	Thr 95	Leu	Glu	Thr	Val	Leu	Glu	Lys	Leu	Glu	
Glu	Leu	Gln	Glu		Asn	Gln	Asn	Gln	100 Gln 115	Ala	Leu	Lys	Gln	105 Ser 120
Phe	Leu	Glu	Leu		Glu	Leu	Lys	Tyr	Leu 130	Leu	Lys	Lys	Thr	Gln 135
Asp	Phe	Phe	Glu		Glu	Thr	Asn	Leu	Ala 145	Asp	Asp	Phe	Phe	
Glu	Asp	Thr	Ser		Leu	Leu	Glu	Leu	Lys 160	Ala	Val	Pro	Ala	
Met	Thr	Gly	Lys		Gly	Phe	Ile	Ala	Gly 175	Cys	Asp	Pro	Thr	Gly 180
Lys	Arg	Met	Ala	Ser 185	Phe	Glu	Arg	Leu	Leu 190	Trp	Arg	Val	Cys	Arg 195
Gly	Asn	Val	Tyr	Leu 200	Lys	Phe	Ser	Glu	Met 205	Asp	Ala	Pro	Leu	Glu 210
Asp	Pro	Val	Thr		Glu	Glu	Ile	Gln	Lys 220	His	Ile	Phe	Ile	Ile 225
Phe	Tyr	Gln	Gly		Gln	Leu	Arg	Gln	Lys 235	Ile	Lys	Lys	Ile	Cys 240
Asp	Gly	Phe	Arg		Thr	Val	Tyr	Pro	Cys 250	Pro	Glu	Pro	Ala	
Glu	Arg	Arg	Glu	Met 260	Leu	Glu	Ser	Val	Asn 265	Val	Arg	Leu	Glu	
Leu	Ile	Thr	Val		Thr	Gln	Thr	Glu	Ser 280	His	Arg	Gln	Arg	
Leu	Gln	Glu	Ala	Ala 290	Ala	Asn	Trp	His	Ser 295	Trp	Leu	Ile	Lys	Val 300
Gln	Lys	Met	Lys	Ala 305	Val	Tyr	His	Ile	Leu 310	Asn	Met	Суѕ	Asn	Ile 315
Asp	Val	Thr	Gln		Cys	Val	Ile	Ala	Glu 325	Ile	Trp	Phe	Pro	
Ala	Asp	Ala	Thr	Arg 335	Ile	Lys	Arg	Ala	Leu 340	Glu	Gln	Gly	Met	
Leu	Ser	Gly	Ser		Met	Ala	Pro	Ile	Met 355	Thr	Thr	Val	Gln	
Lys	Thr	Ala	Pro		Thr	Phe	Asn	Arg	Thr 370	Asn	Lys	Phe	Thr	Ala 375
Gly	Phe	Gln	Asn		Val	Asp	Ala	Tyr	Gly 385	Val	Gly	Ser	Tyr	
Glu	Ile	Asn	Pro		Pro	туr	Thr	Ile	Ile 400	Thr	Phe	Pro	Phe	
Phe	Ala	Val	Met		Gly	Asp	Cys	Gly	His 415	Gly	Thr	Val	Met	Leu 420
Leu	Ala	Ala	Leu		Met	Ile	Leu	Asn	Glu 430	Arg	Arg	Leu	Leu	Ser 435
Gln	Lys	Thr	Asp	Asn 440	Glu	Ile	Trp	Asn	Thr	Phe	Phe	His	Gly	Arg 450
Tyr	Leu	Ile	Leu		Met	Gly	Ile	Phe	Ser 460	Ile	Tyr	Thr	Gly	
Ile	Tyr	Asn			Phe	Ser	Lys	Ser	Leu 475	Asn	Ile	Phe	Gly	Ser 480
Ser	Trp	Ser	Val		Pro	Met	Phe	Arg	Asn 490	Gly	Thr	Trp	Asn	Thr 495
His	Val	Met	Glu		Ser	Leu	Tyr	Leu	Gln 505	Leu	Asp	Pro	Ala	Ile 510
Pro	Gly	Val	Tyr		Gly	Asn	Pro	Tyr	Pro 520	Phe	Gly	Ile	Asp	Pro 525
Ile	Trp	Asn	Leu		Ser	Asn	Lys	Leu	Thr 535	Phe	Leu	Asn	Ser	Tyr
Lys	Met	Lys	Met		Val	Ile	Leu	Gly	Ile	Val	Gln	Met	Val	
Gly	Val	Ile	Leu		Leu	Phe	Asn	His	550 Ile	Tyr	Phe	Arg	Arg	
Leu	Asn	Ile	Ile		Gln	Phe	Ile	Pro	565 Glu	Met	Ile	Phe	Ile	
Cys	Leu	Phe	Gly	Tyr	Leu	Val	Phe	Met	580 Ile	Ile	Phe	Lys	Trp	
				590					595					600

```
Cys Phe Asp Val His Val Ser Gln His Ala Pro Ser Ile Leu Ile
                605
                                    610
His Phe Ile Asn Met Phe Leu Phe Asn Tyr Ser Asp Ser Ser Asn
                620
                                     625
Ala Pro Leu Tyr Lys His Gln Gln Glu Val Gln Ser Phe Phe Val
                635
                                     640
                                                          645
Val Met Ala Leu Ile Ser Val Pro Trp Met Leu Leu Ile Lys Pro
                650
                                    655
                                                          660
Phe Ile Leu Arg Ala Ser His Arg Lys Ser Gln Leu Gln Ala Ser
                 665
                                     670
                                                          675
Arg Ile Gln Glu Asp Ala Thr Glu Asn Ile Glu Gly Asp Ser Ser
                680
                                     685
                                                          690
Ser Pro Ser Ser Arg Ser Gly Gln Arg Thr Ser Ala Asp Thr His
                695
                                     700
                                                          705
Gly Ala Leu Asp Asp His Gly Glu Glu Phe Asn Phe Gly Asp Val
                710
                                     715
                                                          720
Phe Val His Gln Ala Ile His Thr Ile Glu Tyr Cys Leu Gly
                                                          Cys
                725
                                     730
                                                          735
Ile Ser Asn Thr Ala Ser Tyr Leu Arg Leu Trp Ala Leu Ser Leu
                 740
                                     745
                                                          750
Ala His Ala Gln Leu Ser Glu Val Leu Trp Thr Met Val Met Asn
                 755
                                     760
                                                          765
Ser Gly Leu Gln Thr Arg Gly Trp Gly Gly Ile Val Gly Val Phe
                770
                                     775
                                                          780
Ile Ile Phe Ala Val Phe Ala Val Leu Thr Val Ala Ile Leu Leu
                785
                                     790
                                                          795
Ile Met Glu Gly Leu Ser Ala Phe Leu His Ala Leu Arg Leu His
                800
                                     805
                                                          810
Trp Val Glu Phe Gln Asn Lys Phe Tyr Val Gly Asp Gly Tyr Lys
                815
                                                          825
                                     820
Phe Ser Pro Phe Ser Phe Lys His Ile Leu Asp Gly Thr Ala Glu
                830
Glu
<210> 21
<211> 253
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 3715961CD1
<400> 21
Met Ser Glu Cys Pro Leu Ile Leu Tyr Ile His Lys His Ile Asp
                                      10
                                                           15
Thr Tyr Ser Gln Ser Tyr Leu Phe Asn Asp Leu Phe Tyr Pro Val
                  20
                                      25
                                                           30
Tyr Ser Gly Gly Arg Met Val Thr Tyr Glu His Leu Arg Glu Val
                                      40
                                                           45
Val Phe Gly Lys Ser Glu Asp Glu His Tyr Pro Leu Trp Lys Ser
                  50
                                      55
                                                            60
Val Ile Gly Gly Met Met Ala Gly Val Ile Gly Gln Phe Leu Ala
                  65
                                      70
Asn Pro Thr Asp Leu Val Lys Val Gln Met Gln Met Glu Gly Lys
                  80
                                      85
                                                            90
Arg Lys Leu Glu Gly Lys Pro Leu Arg Phe Arg Gly Val His His
                  95
                                     100
                                                          105
Ala Phe Ala Lys Ile Leu Ala Glu Gly Gly Ile Arg Gly Leu Trp
                110
                                     115
                                                          120
Ala Gly Trp Val Pro Asn Ile Gln Arg Ala Ala Leu Val Asn Met
                125
                                     130
                                                          135
Gly Asp Leu Thr Thr Tyr Asp Thr Val Lys His Tyr Leu Val Leu
                                     145
                                                          150
Asn Thr Pro Leu Glu Asp Asn Ile Met Thr His Gly Leu Ser Ser
                155
                                     160
                                                          165
```

Leu Cys Ser Gly Leu Val Ala Ser Ile Leu Gly Thr Pro Ala Asp

```
Val Ile Lys Ser Arg Ile Met Asn Gln Pro Arg Asp Lys Gln Gly
                185
                                      190
                                                           195
Arg Gly Leu Leu Tyr Lys Ser Ser Thr Asp Cys Leu Ile Gln Ala
                200
                                      205
                                                           210
Val Gln Gly Glu Gly Phe Met Ser Leu Tyr Lys Gly Phe Leu Pro
                215
                                      220
                                                           225
Ser Trp Leu Arg Met Thr Pro Trp Ser Met Val Phe Trp Leu Thr
                230
                                      235
Tyr Glu Lys Ile Arg Glu Met Ser Gly Val Ser Pro Phe
                245
                                      250
<210> 22
<211> 229
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5108194CD1
<400> 22
Met Gly Asn Gly Val Lys Glu Gly Pro Val Arg Leu His Glu Asp
                                       10
Ala Glu Ala Val Leu Ser Ser Ser Val Ser Ser Lys Arg Asp His
                 20
                                       25
                                                           30
Arg Gln Val Leu Ser Ser Leu Leu Ser Gly Ala Leu Ala Gly Ala
                 35
                                       40
                                                            45
Leu Ala Lys Thr Ala Val Ala Pro Leu Asp Arg Thr Lys Ile Ile
                 50
                                       55
Phe Gln Val Ser Ser Lys Arg Phe Ser Ala Lys Glu Ala Phe Arg
                 65
                                       70
                                                           75
Val Leu Tyr Tyr Thr
                    Tyr Leu Asn Glu Gly Phe Leu Ser Leu Trp
                 80
                                       85
                                                            90
Arg Gly Asn Ser Ala Thr Met Val Arg Val Val Pro Tyr Ala Ala
                 95
                                      100
                                                          105
Ile Gln Phe Ser Ala His Glu Glu Tyr Lys Arg Ile Leu Gly Ser
                110
                                      115
                                                          120
Tyr Tyr Gly Phe Arg Gly Glu Ala Leu Pro Pro Trp Pro Arg Leu
                125
                                      130
                                                           135
Phe Ala Gly Ala Leu Ala Gly Thr Thr Ala Ala Ser Leu Thr Tyr
                140
                                      145
                                                           150
Pro Leu Asp Leu Val Arg Ala Arg Met Ala Val Thr Pro Lys Glu
                155
                                      160
                                                           165
Met Tyr Ser Asn Ile Phe His Val Phe Ile Arg Ile Ser Arg Glu
                170
                                      175
                                                          180
Glu Gly Leu Lys Thr Leu Tyr His Gly Phe Met Pro Thr Val Leu
                185
                                      190
                                                           195
Gly Val Ile Pro Tyr Ala Gly Leu Ser Phe Phe Thr Tyr Glu Thr
                200
                                      205
                                                          210
Leu Lys Ser Leu His Arg Glu Tyr Ser Gly Arg Lys Leu Ile Pro
                215
                                      220
Phe Ser Glu Gly
<210> 23
<211> 170
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5503122CD1
<400> 23
Met Tyr Asp Asn Leu Tyr Leu His Gly Ile Glu Asp Ser Glu Ala
                                       10
                                                           15
Gly Ser Ala Asp Ser Tyr Thr Ser Arg Pro Ser Asp Ser Asp Val
                 20
                                       25
                                                           30
Ser Leu Glu Glu Asp Arg Glu Ala Ile Arg Gln Glu Arg Glu Gln
                 35
                                       40
```

```
Gln Ala Ala Ile Gln Leu Glu Arg Ala Lys Ser Lys Pro Val Ala
                  50
                                      55
Phe Ala Val Lys Thr Asn Val Ser Tyr Cys Gly Ala Leu Asp Glu
                  65
                                                          75
Asp Val Pro Val Pro Ser Thr Ala Ile Ser Phe Asp Ala Lys
                                                         Asp
                  80
                                      85
                                                          90
Phe Leu His Ile Lys Glu Lys Tyr Asn Asn Asp Trp Trp
                                                         Glv
                  95
                                     100
                                                         105
Arg Leu Val Lys Glu Gly Cys Glu Ile Gly Phe Ile Pro Ser Pro
                110
                                     115
                                                         120
Leu Arg Leu Glu Asn Ile Arg Ile Gln Gln Glu Gln Lys Arg Gly
                 125
                                     130
                                                         135
Arg Phe His Gly Gly Lys Ser Ser Gly Asn Ser Ser Ser Leu
                 140
                                     145
                                                         150
Gly Glu Met Val Ser Gly Thr Phe Arg Ala Thr Pro Thr Ser Thr
                 155
                                     160
Gly Glu Gly Cys Ser
                170
<210> 24
<211> 655
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5517972CD1
<400> 24
Met Ser Ser Asn Val Glu Val Phe Ile Pro Val Ser Gln Gly
                                      10
Asn Thr Asn Gly Phe Pro Ala Thr Ala Ser Asn Asp Leu Lys
                  20
                                      25
                                                          30
Phe Thr Glu Gly Ala Val Leu Ser Phe His Asn Ile Cys
                                                     Tyr Arg
                  35
                                      40
Val Lys Leu Lys Ser Gly Phe Leu Pro Cys Arg Lys Pro Val Glu
                  50
Lys Glu Ile Leu Ser Asn Ile Asn Gly Ile Met Lys Pro Gly Leu
                  65
                                      70
                                                          75
Asn Ala Ile Leu Gly Pro Thr Gly Gly Gly Lys Ser Ser Leu Leu
                 80
                                      85
                                                           90
Asp Val Leu Ala Ala Arg Lys Asp Pro Ser Gly Leu Ser Gly Asp
                  95
                                     100
                                                          105
Val Leu Ile Asn Gly Ala Pro Arg Pro Ala Asn Phe Lys Cys Asn
                 110
                                     115
                                                          120
Ser Gly Tyr Val Val Gln Asp Asp Val Val Met Gly Thr Leu Thr
                 125
                                     130
                                                          135
Val Arg Glu Asn Leu Gln Phe Ser Ala Ala Leu Arg Leu Ala Thr
                 140
                                     145
                                                          150
Thr Met Thr Asn His Glu Lys Asn Glu Arg Ile Asn Arg Val Ile
                 155
                                     160
                                                          165
Gln Glu Leu Gly Leu Asp Lys Val Ala Asp Ser Lys Val Gly Thr
                 170
                                     175
                                                          180
Gln Phe Ile Arg Gly Val Ser Gly Gly Glu Arg Lys Arg Thr Ser
                 185
                                     190
                                                          195
Ile Gly Met Glu Leu Ile Thr Asp Pro Ser Ile Leu Phe Leu Asp
                 200
                                     205
                                                          210
Glu Pro Thr Thr Gly Leu Asp Ser Ser Thr Ala Asn Ala Val Leu
                 215
                                     220
Leu Leu Lys Arg Met Ser Lys Gln Gly Arg Thr Ile Ile Phe
                 230
                                     235
                                                          240
Ser Ile His Gln Pro Arg Tyr Ser Ile Phe Lys Leu Phe Asp Ser
                 245
                                     250
                                                          255
Leu Thr Leu Leu Ala Ser Gly Arg Leu Met Phe His Gly Pro Ala
                 260
                                     265
                                                          270
Gln Glu Ala Leu Gly Tyr Phe Glu Ser Ala Gly Tyr His Cys Glu
                 275
                                     280
                                                          285
Ala Tyr Asn Asn Pro Ala Asp Phe Phe Leu Asp Ile Ile Asn
                                                         Glv
                 290
                                     295
                                                          300
```

```
Asp Ser Thr Ala Val Ala Leu Asn Arg Glu Glu Asp Phe Lys Ala
                305
                                     310
Thr Glu Ile Ile Glu Pro Ser Lys Gln Asp Lys Pro Leu Ile Glu
                320
                                     325
                                                          330
Lys Leu Ala Glu Ile Tyr Val Asn Ser Ser Phe Tyr Lys Glu Thr
                335
                                     340
                                                          345
Lys Ala Glu Leu His Gln Leu Ser Gly Gly Glu Lys Lys
                                                          Lvs
                350
                                     355
Ile Thr Val Phe Lys Glu Ile Ser Tyr Thr Thr Ser Phe Cys His
                365
                                     370
                                                          375
Gln Leu Arg Trp Val Ser Lys Arg Ser Phe Lys Asn Leu Leu Gly
                380
                                     385
                                                          390
Asn Pro Gln Ala Ser Ile Ala Gln Ile Ile Val Thr Val Val Leu
                395
                                     400
                                                          405
Gly Leu Val Ile Gly Ala Ile Tyr Phe Gly Leu Lys Asn Asp Ser
                410
                                     415
                                                          420
Thr Gly Ile Gln Asn Arg Ala Gly Val Leu Phe Phe Leu Thr Thr
                425
                                     430
Asn Gln Cys Phe Ser Ser Val Ser Ala Val Glu Leu Phe Val Val
                440
                                     445
                                                          450
Glu Lys Lys Leu Phe Ile His Glu Tyr Ile Ser Gly Tyr Tyr Arg
                455
                                     460
                                                          465
Val Ser Ser Tyr Phe Leu Gly Lys Leu Leu Ser Asp Leu Leu Pro
                470
                                     475
                                                          480
Met Arg Met Leu Pro Ser Ile Ile Phe Thr Cys Ile Val Tyr Phe
                485
                                     490
                                                          495
Met Leu Gly Leu Lys Pro Lys Ala Asp Ala Phe Phe Val Met Met
                500
                                     505
                                                          510
Phe Thr Leu Met Met Val Ala Tyr Ser Ala Ser Ser Met Ala Leu
                515
                                     520
Ala Ile Ala Ala Gly Gln Ser Val Val Ser Val Ala Thr Leu Leu
                530
                                     535
                                                          540
Met Thr Ile Cys Phe Val Phe Met Met Ile Phe Ser Gly Leu Leu
                545
                                                          555
                                     550
Val Asn Leu Thr Thr Ile Ala Ser Trp Leu Ser Trp Leu Gln Tyr
                560
                                     565
Phe Ser Ile Pro Arg Tyr Gly Phe Thr Ala Leu Gln His Asn Glu
                575
                                     580
                                                          585
Phe Leu Gly Gln Asn Phe Cys Pro Gly Leu Asn Ala Thr Gly Asn
                590
                                     595
                                                          600
Asn Pro Cys Asn Tyr Ala Thr Cys Thr Gly Glu Glu Tyr Leu Val
                605
                                     610
                                                          615
Lys Gln Gly Ile Asp Leu Ser Pro Trp Gly Leu Trp Lys Asn His
                620
                                     625
                                                          630
Val Ala Leu Ala Cys Met Ile Val Ile Phe Leu Thr Ile Ala Tyr
                635
                                     640
Leu Lys Leu Leu Phe Leu Lys Lys Tyr Ser
                650
                                     655
<210> 25
<211> 184
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5593114CD1
<400> 25
Met Trp Val Phe Gly Tyr Gly Ser Leu Ile Trp Lys Val Asp Phe
Pro Tyr Gln Asp Lys Leu Val Gly Tyr Ile Thr Asn Tyr Ser Arg
                 20
                                      25
                                                           30
Arg Phe Trp Gln Gly Ser Thr Asp His Arg Gly Val Pro Gly Lys
                                       40
Pro Gly Arg Val Val Thr Leu Val Glu Asp Pro Ala Gly Cys Val
                 50
                                      55
                                                           60
Trp Gly Val Ala Tyr Arg Leu Pro Val Gly Lys Glu Glu Glu Val
                                      70
```

```
Lys Ala Tyr Leu Asp Phe Arg Glu Lys Gly Gly Tyr Arg Thr Thr
                 80
                                      85
Thr Val Ile Phe Tyr Pro Lys Asp Pro Thr Thr Lys Pro Phe Ser
                  95
                                     100
Val Leu Leu Tyr Ile Gly Thr Cys Asp Asn Pro Asp Tyr Leu Gly
                 110
                                     115
                                                         120
Pro Ala Pro Leu Glu Asp Ile Ala Glu Gln Ile Phe Asn Ala Ala
                125
                                     130
                                                         135
Gly Pro Ser Gly Arg Asn Thr Glu Tyr Leu Phe Glu Leu Ala Asn
                140
                                     145
                                                         150
Ser Ile Arg Asn Leu Val Pro Glu Glu Ala Asp Glu His Leu Phe
                155
                                     160
                                                         165
Ala Leu Glu Lys Leu Val Lys Glu Arg Leu Glu Gly Lys Gln Asn
                170
Leu Asn Cys Ile
<210> 26
<211> 154
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 044775CD1
<400> 26
Met Gly Ala Phe Glu Cys Val Arg Lys Val Tyr Gln Thr Asp Gly
 1
                                      10
                                                           15
Leu Lys Gly Phe Tyr Arg Gly Met Ser Ala Ser Tyr Ala Gly Ile
                  20
                                      25
                                                           30
Ser Glu Thr Val Ile His Phe Val Ile Tyr Glu Ser Ile Lys Gln
                  35
                                      40
                                                           45
Lys Leu Leu Glu Tyr Lys Thr Ala Ser Thr Met Glu Asn Asp Glu
                  50
                                      55
                                                           60
Glu Ser Val Lys Glu Ala Ser Asp Phe Val Gly Met Met Leu Ala
                  65
                                      70
Ala Ala Thr Ser Lys Thr Cys Ala Thr Thr Ile Ala Tyr Pro His
                  80
                                      85
                                                           90
Glu Val Val Arg Thr Arg Leu Arg Glu Glu Gly Thr Lys Tyr Arg
                  95
                                                          105
                                     100
Ser Phe Phe Gln Thr Leu Ser Leu Leu Val Gln Glu Glu Gly Tyr
                 110
                                     115
                                                          120
Gly Ser Leu Tyr Arg Gly Leu Thr Thr His Leu Val Arg Gln Ile
                 125
                                     130
                                                          135
Pro Asn Thr Ala Ile Met Met Ala Thr Tyr Glu Leu Val Val Tyr
                 140
                                     145
                                                          150
Leu Leu Asn Gly
<210> 27
<211> 438
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 116588CD1
<400> 27
Met Leu Leu Val Thr Pro Arg Pro Glu Arg Gly Gly Arg Gly Thr
                                      10
                                                           15
Glu Leu Gly Glu Phe Cys Gly Thr Pro Leu Leu Phe Ser Ser Tyr
                  20
                                      25
Phe Cys Tyr Asp Asn Pro Ala Ala Leu Gln Thr Gln Val Lys Arg
                  35
                                      40
                                                           45
Asp Met Gln Val Asn Thr Thr Lys Phe Met Leu Leu Tyr Ala Trp
                  50
                                      55
Tyr Ser Trp Pro Asn Val Val Leu Cys Phe Phe Gly Gly Phe Leu
                  65
```

```
Ile Asp Arg Val Phe Gly Ile Arg Trp Gly Thr Ile Ile Phe Ser
                80
                                     85
Cys Phe Val Cys Ile Gly Gln Val Val Phe Ala Leu Gly Gly Ile
                 95
                                    100
                                                          105
Phe Asn Ala Phe Trp
                   Leu Met Glu Phe Gly Arg Phe Val Phe Gly
                                    115
               110
                                                          120
Ile Gly Gly Glu Ser Leu Ala Val Ala Gln Asn Thr Tyr Ala Val
                125
                                    130
                                                          135
Ser Trp Phe Lys Gly Lys Glu Leu Asn Leu Val Phe Gly Leu Gln
                                     145
                140
                                                          150
Leu Ser Met Ala Arg Ile Gly Ser Thr Val Asn Met Asn Leu Met
                155
                                     160
                                                          165
Gly Trp Leu Tyr Ser Lys Ile Glu Ala Leu Leu Gly Ser Ala Gly
               170
                                     175
                                                          180
His Thr Thr Leu Gly Ile Thr Leu Met Ile Gly Gly Val Thr Cys
                185
                                     190
                                                          195
Ile Leu Ser Leu Ile Cys Ala Leu Ala Leu Ala Tyr Leu Asp Gln
                200
                                     205
                                                          210
Arg Ala Glu Arg Ile Leu His Lys Glu Gln Gly Lys Thr Gly Glu
                215
                                     220
                                                          225
Val Ile Lys Leu Thr Asp Val Lys Asp Phe Ser Leu Pro Leu Trp
                230
                                     235
                                                          240
Leu Ile Phe Ile Ile Cys Val Cys Tyr
                                     Tyr Val Ala Val Phe Pro
                245
                                     250
                                                          255
Phe Ile Gly Leu Gly Lys Val Phe Phe Thr Glu Lys Phe Gly Phe
                260
                                     265
                                                          270
Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser Val Val Tyr Val Ile
                275
                                     280
                                                          285
Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu Val Asp Lys Thr
                290
                                     295
                                                          300
Gly Lys Asn Ile Ile Trp Val Leu Cys Ala Val Ala Ala Thr Leu
                305
                                     310
                                                          315
Val Ser His Met Met Leu Ala Phe Thr Met Trp Asn Pro Trp Ile
                320
                                     325
                                                          330
Ala Met Cys Leu Leu Gly Leu Ser Tyr Ser Leu Leu Ala Cys Ala
                                     340
                                                          345
                335
Leu Trp Pro Met Val Ala Phe Val Val Pro Glu His Gln Leu Gly
                                     355
                350
                                                          360
Thr Ala Tyr Gly Phe Met Gln Ser Ile Gln Asn Leu Gly Leu Ala
                365
                                     370
                                                          375
Ile Ile Ser Ile Ile Ala Gly Met Ile Leu Asp Ser Arg Gly
                                                          Tyr
                380
                                     385
                                                          390
Leu Phe Leu Glu Val Phe Phe Ile Ala Cys Val Ser Leu Ser Leu
                395
                                     400
Leu Ser Val Val Leu Leu Tyr Leu Val Asn Arg Ala Gln Gly Gly
                410
                                                          420
                                     415
Asn Leu Asn Tyr Ser Ala Arg Gln Arg Glu Glu Ile Lys Phe
                                                         Ser
                425
                                     430
His Thr Glu
<210> 28
<211> 237
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 875369CD1
<400> 28
Met Ala His Val Gly Ser Arg Lys Arg Ser Arg Ser Arg Ser Arg
                                      10
Ser Arg Gly Arg Gly Ser Glu Lys Arg Lys Lys Lys Ser Arg Lys
                 20
                                      25
                                                            30
Asp Thr Ser Arg Asn Cys Ser Ala Ser Thr Ser Gln Gly Arg Lys
                 35
                                      40
                                                            45
Ala Ser Thr Ala Pro Gly Ala Glu Ala Ser Pro Ser Pro Cys Ile
                 50
                                      55
```

```
Thr Glu Arg Ser Lys Gln Lys Ala Arg Arg Arg Thr Arg Ser Ser
                 65
                                      70
Ser Ser Ser Ser
                    Ser Ser Ser Ser Ser Ser Ser Ser
                                                          90
                 80
                                      85
Ser Ser Ser Ser
                    Ser Ser Ser Asp Gly Arg Lys Lys Arg Gly
                 95
                                     100
                                                         105
Lys Tyr Lys Asp Lys Arg Arg Lys Lys Lys Lys Arg Lys Lys
                110
                                                         120
Leu Lys Lys Gly Lys Glu Lys Ala Glu Ala Gln Gln Val Glu
                125
                                     130
                                                         135
Ala Leu Pro Gly Pro Ser Leu Asp Gln Trp His Arg Ser Ala Gly
                140
                                     145
                                                         150
Glu Glu Glu Asp Gly Pro Val Leu Thr Asp Glu Gln Lys Ser
                                                         Arq
                155
                                     160
                                                         165
Ile Gln Ala Met Lys Pro Met Thr Lys Glu Glu Trp Asp Ala Arg
                170
                                     175
                                                         180
Gln Ser Ile Ile Arg Lys Val Val Asp Pro Glu Thr Gly Arg
                                                         Thr
                185
                                     190
                                                         195
Arg Leu Ile Lys Gly Asp Gly Glu Val Leu Glu Glu Ile Val Thr
                200
                                     205
                                                         210
Lys Glu Arg His Arg Glu Ile Asn Lys Gln Ala Thr Arg Gly Asp
                215
                                     220
                                                         225
Cys Leu Ala Phe Gln Met Arg Ala Gly Leu Leu Pro
                230
                                     235
<210> 29
<211> 219
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 1325518CD1
<400> 29
Met Lys Leu Leu Trp Ala Cys Ile Val Cys Val Ala Phe Ala
Arg Lys Arg Arg Phe Pro Phe Ile Gly Glu Asp Asp Asn Asp Asp
                 20
                                      25
                                                          30
Gly His Pro Leu His
                    Pro Ser Leu Asn Ile Pro Tyr Gly Ile Arg
                 35
                                      40
                                                          45
Asn Leu Pro Pro Pro Leu Tyr Tyr Arg Pro Val Asn Thr Val Pro
                 50
                                      55
                                                          60
Ser Tyr Pro Gly Asn Thr Tyr Thr Asp Thr Gly Leu Pro Ser Tyr
                 65
                                      70
Pro Trp Ile Leu Thr Ser Pro Gly Phe Pro Tyr Val Tyr His Ile
                 80
                                      85
Arg Gly Phe Pro Leu Ala Thr Gln Leu Asn Val Pro Pro Leu Pro
                 95
                                     100
                                                         105
Pro Arg Gly Phe Pro Phe Val Pro Pro Ser Arg Phe Phe Ser Ala
                110
                                     115
                                                         120
Ala Ala Pro Ala Ala Pro Pro Ile Ala Ala Glu Pro Ala Ala
                125
                                     130
                                                         135
Ala Ala Pro Leu Thr Ala Thr Pro Val Ala Ala Glu Pro Ala Ala
                140
                                                         150
                                     145
Gly Ala Pro Val Ala Ala Glu Pro Ala Ala Glu Ala Pro Val Gly
                155
                                     160
                                                         165
Ala Glu Pro Ala Ala Glu Ala Pro Val Ala Ala Glu Pro Ala Ala
                170
                                     175
Glu Ala Pro Val Gly Val Glu Pro Ala Ala Glu Glu Pro Ser Pro
                 185
                                     190
                                                         195
Ala Glu Pro Ala Thr Ala Lys Pro Ala Ala Pro Glu Pro His Pro
                200
                                     205
                                                         210
Ser Pro Ser Leu Glu Gln Ala Asn Gln
                 215
<210> 30
<211> 707
<212> PRT
<213> Homo sapiens
```

<220>

<221> misc_feature <223> Incyte ID No: 2060987CD1 <400> 30 Met Ala Ala Ala Thr Ala Ala Glu Gly Val Pro Ser Arg Gly Pro Gly Glu Val Ile His Leu Asn Val Gly Gly Lys Arg Phe Ser Thr Ser Arg Gln Thr Leu Thr Trp Ile Pro Asp Ser Phe Phe Ser Ser Leu Leu Ser Gly Arg Ile Ser Thr Leu Lys Asp Glu Thr Gly Ala Ile Phe Ile Asp Arg Asp Pro Thr Val Phe Ala Pro Ile Leu Asn Phe Leu Arg Thr Lys Glu Leu Asp Pro Arg Gly Val His Gly Ser Ser Leu Leu His Glu Ala Gln Phe Tyr Gly Leu Thr Pro Leu Val Arg Arg Leu Gln Leu Arg Glu Glu Leu Asp Arg Ser Ser Cys Gly Asn Val Leu Phe Asn Gly Tyr Leu Pro Pro Pro Val Phe Pro Val Lys Arg Arg Asn Arg His Ser Leu Val Gly Pro Gln Gln Leu Gly Gly Arg Pro Ala Pro Val Arg Arg Ser Asn Thr Met Pro Pro Asn Leu Gly Asn Ala Gly Leu Leu Gly Arg Met Leu Asp Glu Lys Thr Pro Pro Ser Pro Ser Gly Gln Pro Glu Glu Pro Gly Met Arg Leu Val Cys Gly His His Asn Trp Ile Ala Val Ala Tyr Thr Gln Phe Leu Val Cys Tyr Arg Leu Lys Glu Ala Ser Gly Trp Gln Leu Val Phe Ser Ser Pro Arg Leu Asp Trp Pro Ile Glu Arg Leu Ala Leu Thr Ala Arg Val His Gly Gly Ala Leu Gly Glu His Asp Lys Met Val Ala Ala Thr Gly Ser Glu Ile Leu Leu Trp Ala Leu Gln Ala Glu Gly Gly Ser Glu Ile Gly Val Phe His Leu Gly Val Pro Val Glu Ala Leu Phe Phe Val Gly Asn Gln Leu Ile Ala Thr Ser His Thr Gly Arg Ile Gly Val Trp Asn Ala Val Thr Lys His Trp Gln Val Gln Glu Val Gln Pro Ile Thr Ser Tyr Asp Ala Ala Gly Ser Phe Leu Leu Gly Cys Asn Asn Gly Ser Ile Tyr Tyr Val Asp Val Gln Lys Phe Pro Leu Arg Met Lys Asp Asn Asp Leu Leu Val Ser Glu Leu Tyr Arg Asp Pro Ala Glu Asp Gly Val Thr Ala Leu Ser Val Tyr Leu Thr Pro Lys Thr Ser Asp Ser Gly Asn Trp Ile Glu Ile Ala Tyr Gly Thr Ser Ser Gly Gly Val Arg Val Ile Val Gln His Pro Glu Thr Val Gly Ser Gly Pro Gln Leu Phe Gln Thr Phe Thr Val His Arg Ser Pro Val Thr Lys Ile Met Leu Ser Glu Lys His Leu Ile Ser Val Cys Ala Asp Asn Asn His Val Arg Thr Trp Ser Val Thr Arg Phe Arg Gly Met Ile Ser Thr Gln Pro Gly Ser Thr Pro Leu Ala Ser Phe Lys Ile Leu

```
470
                                      475
                                                          480
Ala Leu Glu Ser Ala Asp Gly His Gly Gly Cys Ser Ala Gly Asn
                 485
                                      490
                                                          495
Asp Ile Gly Pro Tyr Gly Glu Arg Asp Asp Gln Gln Val Phe Ile
                 500
                                      505
Gln Lys Val Val Pro Ser Ala Ser Gln Leu Phe Val Arg Leu Ser
                 515
                                      520
                                                           525
Ser Thr Gly Gln Arg
                     Val Cys Ser Val Arg Ser Val Asp Gly Ser
                 530
                                                           540
                                      535
Pro Thr Thr Ala Phe Thr Val Leu Glu Cys Glu Gly Ser Arg Arg
                 545
                                      550
                                                           555
Leu Gly Ser Arg Pro Arg Arg Tyr Leu Leu Thr Gly Gln Ala Asn
                 560
                                      565
                                                           570
Gly Ser Leu Ala Met Trp Asp Leu Thr Thr Ala Met Asp Gly Leu
                 575
                                      580
                                                           585
Gly Gln Ala Pro Ala Gly Gly Leu Thr Glu Gln Glu Leu Met Glu
                 590
                                      595
                                                           600
Gln Leu Glu His Cys
                     Glu Leu Ala Pro Pro Ala Pro Ser Ala Pro
                 605
                                      610
                                                           615
Ser Trp Gly Cys Leu Pro Ser Pro Ser Pro Arg Ile Ser Leu Thr
                 620
                                      625
                                                           630
Ser Leu His Ser Ala Ser Ser Asn Thr Ser Leu Ser Gly His Arg
                 635
                                      640
                                                           645
Gly Ser Pro Ser Pro
                    Pro Gln Ala Glu Ala Arg Arg Arg Gly Gly
                 650
                                      655
Gly Ser Phe Val Glu Arg Cys Gln Glu Leu Val Arg Ser Gly Pro
                 665
                                      670
                                                           675
Asp Leu Arg Arg Pro Pro Thr Pro Ala Pro
                                          Trp Pro Ser Ser
                                                          Gly
                 680
                                      685
                                                           690
Leu Gly Thr Pro Leu Thr Pro Pro Lys Met Lys Leu Asn Glu Thr
                 695
                                      700
                                                           705
Ser Phe
<210> 31
<211> 279
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2172064CD1
<400> 31
Met Cys Gly Arg Phe Leu Arg Arg Leu Leu Ala Glu Glu Ser Arg
  1
                                       10
                                                            15
Arg Ser Thr Pro Val Gly Arg Leu Leu Pro Val Leu Leu Gly
                  20
                                        25
                                                            30
Phe Arg Leu Val Leu Leu Ala Ala Ser Gly
                                          Pro Gly Val Tyr Gly
                                                            45
                  35
                                        40
Asp Glu Gln Ser Glu Phe Val Cys His Thr Gln Gln Pro Gly Cys
                  50
                                                            60
                                        55
Lys Ala Ala Cys Phe Asp Ala Phe His Pro Leu Ser Pro Leu Arg
                  65
                                        70
Ser Trp Val Phe Gln Val Ile Leu Val Ala Val Pro Ser Ala Leu
                  80
                                        85
                                                            90
Tyr Met Gly Phe Thr Leu Tyr His Val Ile Trp His Trp Glu Leu
                  95
                                       100
                                                           105
Ser Gly Lys Gly Lys
                     Glu Glu Glu Thr Leu Ile Gln Gly Arg Glu
                 110
                                       115
                                                           120
Gly Asn Thr Asp Val Pro Gly Ala Gly Ser Leu Arg Leu Leu Trp
                 125
                                       130
                                                           135
Ala Tyr Val Ala Gln Leu Gly Ala Arg Leu Val Leu Glu Gly Ala
                 140
                                       145
                                                           150
Ala Leu Gly Leu Gln Tyr His Leu Tyr Gly Phe Gln Met Pro Ser
                 155
                                       160
                                                           165
Ser Phe Ala Cys Arg Arg Glu Pro Cys Leu Gly Ser Ile Thr Cys
                  170
                                       175
                                                           180
Asn Leu Ser Arg Pro Ser Glu Lys Thr Ile Phe Leu Lys Thr Met
```

```
185
                                    190
                                                         195
Phe Gly Val Ser Gly Phe Cys Leu Leu Phe Thr Phe Leu Glu Leu
                200
                                    205
                                                          210
Val Leu Leu Gly Leu Gly Arg Trp Trp Arg Thr Trp Lys His Lys
                215
                                    220
                                                          225
Ser Ser Ser Lys Tyr Phe Leu Thr Ser Glu Ser Thr Arg Arg
                230
                                    235
                                                          240
His Lys Lys Ala Thr Asp Ser Leu Pro Val Val Glu Thr Lys Glu
                245
                                    250
                                                          255
Gln Phe Gln Glu Ala Val Pro Gly Arg Ser Leu Ala Gln Glu Lys
                260
                                    265
                                                          270
Gln Arg Pro Val Gly Pro Arg Asp Ala
                275
<210> 32
<211> 154
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2219267CD1
<400> 32
Met Val Thr Gly Leu Ala Ser Leu Leu Leu Ala Gly Ala Gln
                  5
                                     10
                                                           15
Tyr Leu Pro Gly Trp Thr Val Leu Phe Leu Ser Val Leu Gly Leu
                                     25
                                                           30
Leu Ala Ser Arg Ala Val Ser Ala Leu Ser Ser Leu Phe Ala Ala
                 35
                                     40
                                                           45
Glu Val Phe Pro Thr Val Ile Arg Gly Ala Gly Leu Gly Leu Val
                 50
                                     55
                                                           60
Leu Gly Ala Gly Phe Leu Gly Gln Ala Ala Gly Pro Leu Asp Thr
                 65
                                                           75
                                     70
Leu His Gly Arg Gln Gly Phe Phe Leu Gln Gln Val Val Phe Ala
                 80
                                     85
                                                           90
Ser Leu Ala Val Leu Ala Leu Leu Cys Val Leu Leu Leu Pro Glu
                 95
                                     100
Ser Arg Ser Arg Gly Leu Pro Gln Ser Leu Gln Asp Ala Asp Arg
                110
                                     115
                                                          120
Leu Arg Arg Ser Pro Leu Leu Arg Gly Arg Pro Arg Gln Asp His
                125
                                     130
                                                          135
Leu Pro Leu Leu Pro Pro Ser Asn Ser Tyr Trp Ala Gly His Thr
                140
Pro Glu Gln His
<210> 33
<211> 289
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2308629CD1
<400> 33
Met Val Ala Gly Ala Val Ala Gly Ile Leu Glu His Cys Val Met
                                                           15
                                      10
Tyr Pro Ile Asp Cys Val Lys Thr Arg Met Gln Ser Leu Gln Pro
                 20
                                      25
                                                           30
Asp Pro Ala Ala Arg Tyr Arg Asn Val Leu Glu Ala Leu Trp Arg
                 35
                                      40
                                                           45
Ile Ile Arg Thr Glu Gly Leu Trp Arg Pro Met Arg Gly Leu Asn
                 50
                                      55
                                                           60
Val Thr Ala Thr Gly Ala Gly Pro Ala His Ala Leu Tyr Phe Ala
                 65
                                      70
                                                           75
Cys Tyr Glu Lys Leu Lys Lys Thr Leu Ser Asp Val Ile His
                                                         Pro
                 80
                                      85
Gly Gly Asn Ser His Ile Ala Asn Gly Ala Ala Gly Cys Val Ala
```

```
95
                                     100
                                                          105
Thr Leu Leu His Asp Ala Ala Met Asn Pro Ala Glu Val Val Lys
                110
                                     115
                                                          120
Gln Arg Met Gln Met Tyr Asn Ser Pro Tyr His Arg Val Thr Asp
                125
                                     130
Cys Val Arg Ala Val Trp Gln Asn Glu Gly Ala Gly Ala Phe Tyr
                140
                                     145
                                                          150
Arg Ser Tyr Thr Thr Gln Leu Thr Met Asn Val Pro Phe Gln Ala
                155
                                     160
                                                          165
Ile His Phe Met Thr Tyr Glu Phe Leu Gln Glu His Phe Asn Pro
                170
                                     175
                                                          180
Gln Arg Arg Tyr Asn Pro Ser Ser His Val Leu Ser Gly Ala Cys
                185
                                     190
                                                          195
Ala Gly Ala Val Ala Ala Ala Thr Thr
                                         Pro Leu Asp Val Cys
                                                          210
                200
                                     205
Lys Thr Leu Leu Asn Thr Gln Glu Ser Leu Ala Leu Asn Ser His
                215
                                                          225
                                     220
Ile Thr Gly His Ile
                    Thr Gly Met Ala Ser Ala Phe Arg Thr Val
                230
                                     235
                                                          240
Tyr Gln Val Gly Gly Val Thr Ala Tyr Phe Arg Gly Val Gln Ala
                245
                                                          255
                                     250
Arg Val Ile Tyr Gln Ile Pro Ser Thr Ala Ile Ala Trp Ser Val
                260
                                     265
Tyr Glu Phe Phe Lys Tyr Leu Ile Thr Lys Arg Gln Glu Glu Trp
                                      280
Arg Ala Gly Lys
```

<210> 34

<211> 300

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2660038CD1

<400> 34 Met Asp Phe Leu Met Ser Gly Leu Ala Ala Cys Gly Ala Cys Val Phe Thr Asn Pro Leu Glu Val Val Lys Thr Arg Met Gln Leu Gln Gly Glu Leu Gln Ala Pro Gly Thr Tyr Gln Arg His Tyr Arg Asn Val Phe His Ala Phe Ile Thr Ile Gly Lys Val Asp Gly Leu Ala Ala Leu Gln Lys Gly Leu Ala Pro Ala Leu Leu Tyr Gln Phe Leu Met Asn Gly Ile Arg Leu Gly Thr Tyr Gly Leu Ala Glu Ala Gly Gly Tyr Leu His Thr Ala Glu Ala Thr His Ser Pro Ala Arg Ser Ala Ala Gly Ala Met Ala Gly Val Met Gly Ala Tyr Leu Gly Ser Pro Ile Tyr Met Val Lys Thr His Leu Gln Ala Gln Ala Ala Ser Glu Ile Ala Val Gly His Gln Tyr Lys His Gln Gly Met Phe Gln Ala Leu Thr Glu Ile Gly Gln Lys His Gly Leu Val Gly Leu Trp Arg Gly Ala Leu Gly Gly Leu Pro Arg Val Ile Val Gly Ser Ser Thr Gln Leu Cys Thr Phe Ser Ser Thr Lys Asp Leu Leu Ser Gln Trp Glu Ile Phe Pro Pro Gln Ser Trp Lys Leu Ala Leu Val Ala Ala Met Met Ser Gly Ile Ala Val Val Leu Ala Met Ala Pro Phe Asp Val Ala Cys Thr Arg Leu Tyr Asn Gln Pro Thr Asp Ala

```
230
                                                          240
Gln Gly Lys Gly Leu Met Tyr Arg Gly Ile Leu Asp Ala Leu Leu
                 245
                                     250
Gln Thr Ala Arg Thr Glu Gly Ile Phe Gly Met Tyr Lys Gly Ile
                260
                                     265
                                                          270
Gly Ala Ser Tyr Phe Arg Leu Gly Pro His Thr Ile Leu Ser Leu
                275
                                     280
                                                          285
Phe Phe Trp Asp Gln Leu Arg Ser Leu Tyr Tyr Thr Asp Thr
                                                         Lys
                                     295
                                                          300
<210> 35
<211> 382
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2670745CD1
<400> 35
Met Leu Arg Trp Thr Val His Leu Glu Gly Gly Pro Arg Arg Val
                                      10
                                                           15
                                         Val Tyr Ser Phe Gly
Asn His Ala Ala Val Ala Val Gly His Arg
                  20
                                                           30
                                      25
Gly Tyr Cys Ser Gly Glu Asp Tyr Glu Thr Leu Arg Gln Ile Asp
                 35
                                      40
Val His Ile Phe Asn Ala Val Ser Leu Arg Trp Thr Lys Leu Pro
                 50
                                      55
                                                           60
Pro Val Lys Ser Ala Ile Arg Gly Gln Ala Pro Val Val Pro Tyr
                  65
                                      70
                                                           75
Met Arg Tyr Gly His Ser Thr Val Leu Ile Asp Asp Thr Val Leu
                 80
                                      85
                                                           90
Leu Trp Gly Gly Arg Asn Asp Thr Glu Gly Ala Cys Asn Val
                                                          Leu
                 95
                                     100
                                                          105
Tyr Ala Phe Asp Val Asn Thr His Lys Trp Phe Thr Pro Arg Val
                110
                                     115
                                                          120
Ser Gly Thr Val Pro Gly Ala Arg Asp Gly His Ser Ala Cys Val
                125
                                     130
                                                          135
Leu Gly Lys Ile Met Tyr Ile Phe Gly Gly
                                         Tyr Glu Gln Gln Ala
                140
                                     145
                                                          150
Asp Cys Phe Ser Asn Asp Ile His Lys Leu Asp Thr Ser Thr Met
                 155
                                     160
                                                          165
Thr Trp Thr Leu Ile Cys Thr Lys Gly Ser Pro Ala Arg Trp Arg
                170
                                     175
                                                          180
Asp Phe His Ser Ala Thr Met Leu Gly Ser His Met Tyr Val Phe
                 185
                                     190
                                                          195
Gly Gly Arg Ala Asp Arg Phe Gly Pro Phe His Ser Asn Asn Glu
                200
                                     205
                                                          210
Ile Tyr Cys Asn Arg Ile Arg Val Phe Asp Thr Arg Thr Glu Ala
                215
                                     220
                                                          225
            Cys Pro Pro Thr Pro Val Leu Pro Glu Gly Arg Arg
                230
                                     235
                                                          240
Ser His Ser Ala Phe Gly Tyr Asn Gly Glu Leu Tyr Ile Phe Gly
                245
                                     250
                                                          255
Gly Tyr Asn Ala Arg Leu Asn Arg His Phe His Asp Leu Trp
                                                          Lys
                 260
                                     265
                                                          270
Phe Asn Pro Val Ser
                    Phe Thr Trp Lys Lys Ile Glu Pro Lys Gly
                 275
                                     280
                                                          285
Lys Gly Pro Cys Pro Arg Arg Gln Cys Cys Cys Ile Val Gly
                 290
                                     295
                                                           300
Asp Lys Ile Val Leu Phe Gly Gly Thr Ser Pro Ser Pro Glu Glu
                 305
                                     310
                                                          315
Gly Leu Gly Asp Glu Phe Asp Leu Ile Asp His Ser Asp Leu His
                320
                                     325
                                                          330
Ile Leu Asp Phe Ser Pro Ser Leu Lys Thr Leu Cys Lys Leu Ala
                 335
                                     340
                                                          345
Val Ile Gln Tyr Asn Leu Asp Gln Ser Cys Leu Pro His Asp
                                                          Ile
                                     355
                                                          360
```

```
Arg Trp Glu Leu Asn Ala Met Thr Thr Asn Ser Asn Ile Ser Arg
                365
                                     370
Pro Ile Val Ser Ser His Gly
                380
<210> 36 <211> 287
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2676443CD1
<400> 36
Met Ala Ala Glu Ala Arg Val Ser Arg Trp Tyr Phe Gly Gly Leu
 1
Ala Ser Cys Gly Ala Ala Cys Cys Thr His Pro Leu Asp Leu Leu
                 20
                                      25
                                                           30
Lys Val His Leu Gln Thr Gln Gln Glu Val Lys Leu Arg Met Thr
                 35
                                      40
                                                           45
Gly Met Ala Leu Arg Val Val Arg Thr Asp Gly Ile Leu Ala Leu
                 50
                                      55
                                                           60
Tyr Ser Gly Leu Ser Ala Ser Leu Cys Arg Gln Met Thr Tyr Ser
                 65
                                      70
                                                           75
Leu Thr Arg Phe Ala Ile Tyr Glu Thr Val Arg Asp Arg Val Ala
                 80
                                      85
Lys Gly Ser Gln Gly Pro Leu Pro Phe His Glu Lys Val Leu Leu
                 95
                                     100
                                                          105
Gly Ser Val Ser Gly Leu Ala Gly Gly Phe Val Gly Thr Pro Ala
                110
                                     115
                                                          120
Asp Leu Val Asn Val Arg Met Gln Asn Asp Val Lys Leu Pro Gln
                125
                                     130
                                                          135
Gly Gln Arg Arg Asn Tyr Ala His Ala Leu Asp Gly Leu Tyr Arg
                140
                                     145
                                                          150
Val Ala Arg Glu Glu Gly Leu Arg Arg Leu Phe Ser Gly Ala Thr
                155
                                      160
                                                          165
Met Ala Ser Ser Arg Gly Ala Leu Val Thr Val Gly Gln Leu Ser
                170
                                      175
                                                          180
Cys Tyr Asp Gln Ala Lys Gln Leu Val Leu Ser Thr Gly Tyr Leu
                185
                                      190
                                                          195
Ser Asp Asn Ile Phe Thr His Phe Val Ala Ser Phe Ile Ala Gly
                200
                                      205
                                                          210
Gly Cys Ala Thr Phe Leu Cys Gln Pro Leu Asp Val Leu Lys Thr
                215
                                      220
                                                          225
Arg Leu Met Asn Ser Lys Gly Glu Tyr Gln Gly Val Phe His Cys
                230
                                      235
                                                          240
Ala Val Glu Thr Ala Lys Leu Gly Pro Leu Ala Phe Tyr Lys Gly
                245
                                      250
                                                          255
Leu Val Pro Ala Gly Ile Arg Leu Ile Pro His Thr Val Leu Thr
                260
                                     265
                                                          270
Phe Val Phe Leu Glu Gln Leu Arg Lys Asn Phe Gly Ile Lys Val
                275
                                      280
                                                          285
Pro Ser
<210> 37
<211> 497
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 3295764CD1
<400> 37
Met Asp Val Pro Gly Pro Val Ser Arg Arg Ala Ala Ala Ala Ala
                                      10
                                                           15
Ala Thr Val Leu Leu Arg Thr Ala Arg Val Arg Arg Glu Cys Trp
                                       25
```

```
Phe Leu Pro Thr Ala Leu Leu Cys Ala Tyr Gly Phe Phe Ala Ser
                 35
Leu Arg Pro Ser Glu Pro Phe Leu Thr Pro Tyr Leu Leu Gly Pro
                 50
                                     55
                                                           60
Asp Lys Asn Leu Thr Glu Arg Glu Val Phe Asn Glu Ile Tyr Pro
                 65
                                     70
Val Trp Thr Tyr Ser Tyr Leu Val Leu Leu Phe Pro Val Phe Leu
                 80
                                     85
                                                           90
Ala Thr Asp Tyr Leu Arg Tyr Lys Pro Val Val Leu Leu Gln Gly
                 95
                                     100
                                                         105
Leu Ser Leu Ile Val Thr Trp Phe Met Leu Leu Tyr Ala Gln Gly
                110
                                                          120
                                     115
Leu Leu Ala Ile Gln Phe Leu Glu Phe Phe Tyr Gly Ile Ala Thr
                125
                                     130
                                                          135
Ala Thr Glu Ile Ala Tyr Tyr Ser Tyr Ile Tyr Ser Val Val Asp
                140
                                     145
                                                          150
Leu Gly Met Tyr Gln Lys Val Thr Ser Tyr Cys Arg Ser Ala Thr
                155
                                     160
                                                          165
Leu Val Gly Phe Thr Val Gly Ser Val Leu Gly Gln Ile Leu Val
                170
                                     175
                                                          180
Ser Val Ala Gly Trp Ser Leu Phe Ser Leu Asn Val Ile Ser Leu
                185
                                     190
                                                          195
Thr Cys Val Ser Val Ala Phe Ala Val Ala Trp Phe Leu Pro Met
                200
                                     205
                                                          210
Pro Gln Lys Ser Leu Phe Phe His His Ile Pro Ser Thr Cys Gln
                215
                                     220
                                                          225
Arg Val Asn Gly Ile Lys Val Gln Asn Gly Gly Ile Val Thr Asp
                230
                                     235
                                                          240
Thr Pro Ala Ser Asn His Leu Pro Gly Trp Glu Asp Ile Glu Ser
                245
                                     250
                                                          255
Lys Ile Pro Leu Asn Met Glu Glu Pro Pro Val Glu Glu Pro Glu
                260
                                                          270
                                     265
Pro Lys Pro Asp Arg Leu Leu Val Leu Lys Val Leu Trp Asn Asp
                 275
                                     280
                                                          285
Phe Leu Met Cys Tyr Ser Ser Arg Pro Leu Leu Cys Trp Ser Val
                 290
                                     295
                                                          300
Trp Trp Ala Leu Ser Thr Cys Gly Tyr Phe Gln Val Val Asn Tyr
                 305
                                     310
                                                          315
Thr Gln Gly Leu Trp Glu Lys Val Met Pro Ser Arg Tyr Ala Ala
                 320
                                                          330
                                     325
Ile Tyr Asn Gly Gly Val Glu Ala Val Ser Thr Leu Leu Gly Ala
                 335
                                                          345
                                     340
Val Ala Val Phe Ala Val Gly Tyr Ile Lys Ile Ser Trp Ser Thr
                 350
                                     355
Trp Gly Glu Met Thr Leu Ser Leu Phe Ser Leu Leu Ile Ala Ala
                 365
                                     370
                                                          375
Ala Val Tyr Ile Met Asp Thr Val Gly Asn Ile Trp Val Cys Tyr
                 380
                                     385
                                                          390
Ala Ser Tyr Val Val Phe Arg Ile Ile Tyr Met Leu Leu Ile Thr
                                                          405
                 395
                                     400
Ile Ala Thr Phe Gln Ile Ala Ala Asn Leu Ser Met Glu Arg Tyr
                 410
                                                          420
                                     415
Ala Leu Val Phe Gly Val Asn Thr Phe Ile Ala Leu Ala Leu Gln
                 425
                                     430
                                                          435
Thr Leu Leu Thr Leu Ile Val Val Asp Ala Ser Gly Leu Gly Leu
                 440
                                     445
                                                          450
Glu Ile Thr Thr Gln Phe Leu Ile Tyr Ala Ser Tyr Phe Ala Leu
                 455
Ile Ala Val Val Phe Leu Ala Ser Gly Ala Val Ser Val Met Lys
                 470
                                     475
                                                          480
Lys Cys Arg Lys Leu Glu Asp Pro Gln Ser Ser Gln Val Thr
                 485
                                     490
                                                          495
Thr Ser
```

<210> 38

<211> 228

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3438320CD1 Met Pro Arg Arg Gly Leu Val Ala Gly Pro Asp Leu Glu Tyr Phe 10 15 Gln Arg Arg Tyr Phe Thr Pro Ala Glu Val Ala Gln His Asn Arg 20 25 30 Pro Glu Asp Leu Trp Val Ser Tyr Leu Gly Arg Val Tyr Asp Leu 35 40 Thr Ser Leu Ala Gln Glu Tyr Lys Gly Asn Leu Leu Leu Lys Pro 50 55 60 Ile Val Glu Val Ala Gly Gln Asp Ile Ser His Trp Phe Asp Pro 65 70 75 Lys Thr Arg Asp Ile Arg Lys His Ile Asp Pro Leu Thr Gly Cys 80 85 90 Leu Arg Tyr Cys Thr Pro Arg Gly Arg Phe Val His Val Pro Pro 95 105 100 Gln Leu Pro Cys Ser Asp Trp Ala Asn Asp Phe Gly Lys Pro Trp 110 115 120 Trp Gln Gly Ser Tyr Tyr Glu Val Gly Arg Leu Ser Ala Lys Thr 125 130 135 Arg Ser Ile Arg Ile Ile Asn Thr Leu Thr Ser Gln Glu His Thr 140 145 150 Leu Glu Val Gly Val Leu Glu Ser Ile Trp Glu Ile Leu His Arg 155 160 1.65 Tyr Leu Pro Tyr Asn Ser His Ala Ala Ser Tyr Thr Trp Lys Tyr 170 175 180 Glu Gly Lys Asn Leu Asn Met Asp Phe Thr Leu Glu Glu Asn Gly 185 190 195 Ile Arg Asp Glu Glu Glu Phe Asp Tyr Leu Ser Met Asp Gly 200 205 210 Thr Leu His Thr Pro Ala Ile Leu Leu Tyr Phe Asn Asp Asp Leu 215 220 225 Thr Glu Leu <210> 39 <211> 273 <212> PRT <213> Homo sapiens <220> <221> misc_feature <223> Incyte ID No: 3986488CD1 <400> 39 Met Ala Ala Thr Ile Met Ile Leu Tyr Val Ser Lys Leu Asn Lys 10 Ile Ile His Phe Pro Asp Phe Asp Lys Lys Ile Pro Val Lys Leu 20 Phe Pro Leu Pro Leu Leu Tyr Val Gly Asn His Ile Ser Gly Leu 35 45 40 Ser Ser Thr Ser Lys Leu Ser Leu Pro Met Phe Thr Val Leu Arg 50 55 60 Lys Phe Thr Ile Pro Leu Thr Leu Leu Leu Glu Thr Ile Ile Leu 65 75 70 Gly Lys Gln Tyr Ser Leu Asn Ile Ile Leu Ser Val Phe Ala Ile 80 85 Ile Leu Gly Ala Phe Ile Ala Ala Gly Ser Asp Leu Ala Phe Asn 95 100 105 Leu Glu Gly Tyr Ile Phe Val Phe Leu Asn Asp Ile Phe Thr Ala 110 115 120 Ala Asn Gly Val Tyr Thr Lys Gln Lys Met Asp Pro Lys Glu Leu 125 130 Gly Lys Tyr Gly Val Leu Phe Tyr Asn Ala Cys Phe Met Ile Ile 140 145 150 Pro Thr Leu Ile Ile Ser Val Ser Thr Gly Asp Leu Gln Gln Ala

```
155
                                     160
Thr Glu Phe Asn Gln Trp Lys Asn Val Val Phe Ile Leu Gln Phe
                 170
                                     175
                                                          180
Leu Leu Ser Cys Phe Leu Gly Phe Leu Leu Met Tyr Ser Thr
                                                          Val
                 185
                                     190
                                                          195
Leu Cys Ser Tyr Tyr Asn Ser Ala Leu Thr Thr Ala Val Val Gly
                 200
                                     205
                                                          210
Ala Ile Lys Asn Val Ser Val Ala Tyr Ile Gly Ile Leu Ile Gly
                 215
                                     220
                                                          225
Gly Asp Tyr Ile Phe Ser Leu Leu Asn Phe Val Gly Leu Asn Ile
                 230
                                     235
                                                          240
Cys Met Ala Gly Gly Leu Arg Tyr Ser Phe Leu Thr Leu Ser
                 245
                                     250
                                                          255
Gln Leu Lys Pro Lys Pro Val Gly Glu Glu Asn Ile Cys Leu Asp
                 260
                                     265
Leu Lys Ser
<210> 40
<211> 206
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 4378816CD1
<400> 40
Met Gly Ala Glu Trp Glu Leu Gly Ala Glu Ala Gly Gly Ser Leu
                                      10
                                                           15
Leu Leu Cys Ala Ala Leu Leu Ala Ala Gly Cys Ala Leu Gly Leu
                  20
                                      25
                                                           30
Arg Leu Gly Arg Gly Gln Gly Ala Ala Asp Arg Gly Ala Leu Ile
                  35
                                      40
Trp Leu Cys Tyr Asp Ala Leu Val His Phe Ala Leu Glu Gly Pro
                  50
                                      55
Phe Val Tyr Leu Ser Leu Val Gly Asn Val Ala Asn Ser Asp Gly
                  65
                                      70
                                                           75
Leu Ile Ala Ser Leu Trp Lys Glu Tyr Gly Lys Ala Asp Ala Arg
                  80
                                      85
                                                           90
Trp Val Tyr Phe Asp Pro Thr Ile Val Ser Val Glu Ile Leu Thr
                  95
                                     100
                                                          105
Val Ala Leu Asp Gly Ser Leu Ala Leu Phe Leu Ile Tyr Ala Ile
                 110
                                     115
                                                          120
Val Lys Glu Lys Tyr Tyr Arg His Phe Leu Gln Ile Thr Leu Cys
                 125
                                     130
                                                          135
Val Cys Glu Leu Tyr Gly Cys Trp Met Thr Phe Leu Pro Glu Trp
                 140
                                     145
                                                          150
Leu Thr Arg Ser Pro Asn Leu Asn Thr Ser Asn Trp Leu Tyr Cys
                 155
                                     160
                                                          165
Trp Leu Tyr Leu Phe Phe Phe Asn Gly Val Trp Val Leu Ile Pro
                 170
                                     175
                                                          180
Gly Leu Leu Trp Gln Ser Trp Leu Glu Leu Lys Lys Met His
                 185
                                     190
                                                          195
Gln Lys Glu Thr Ser Ser Val Lys Lys Phe Gln
                 200
<210> 41
<211> 235
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 4797137CD1
<400> 41
Met Gln Gln Arg Gly Ala Ala Gly Ser Arg Gly Cys Ala Leu Phe
                                      10
Pro Leu Leu Gly Val Leu Phe Phe Gln Gly Val Tyr Ile Val Phe
```

```
20
                                                            30
Ser Leu Glu Ile Arg Ala Asp Ala His Val Arg Gly Tyr Val Gly
                 35
                                                            45
Glu Lys Ile Lys Leu Lys Cys Thr Phe Lys Ser Thr Ser Asp Val
                 50
                                      55
                                                            60
Thr Asp Lys Leu Thr Ile Asp Trp Thr Tyr Arg Pro Pro Ser Ser
                 65
                                      70
Ser His Thr Val Ser Ile Phe His Tyr Gln Ser Phe Gln Tyr
                                                          Pro
                 80
                                      85
                                                            90
Thr Thr Ala Gly Thr Phe Arg Asp Arg Ile Ser Trp Val Gly Asn
                 95
                                     100
                                                          105
Val Tyr Lys Gly Asp Ala Ser Ile Ser Ile Ser Asn Pro Thr Ile
                110
                                     115
                                                          120
Lys Asp Asn Gly Thr Phe Ser Cys Ala Val Lys Asn Pro Pro Asp
                125
                                     130
                                                          135
Val His His Asn Ile Pro Met Thr Glu Leu Thr Val Thr Glu Arg
                140
                                     145
                                                          150
Gly Phe Gly Thr Met Leu Ser Ser Val Ala Leu Leu Ser Ile Leu
                155
                                     160
                                                          165
Val Phe Val Pro Ser Ala Val Val Val Ala Leu Leu Leu Val Arg
                170
                                     175
                                                          180
Met Gly Arg Lys Ala Ala Gly Leu Lys Lys Arg Ser Arg Ser Gly
                185
                                     190
                                                          195
Tyr Lys Lys Ser Ser Ile Glu Val Ser Asp Asp Thr Asp Gln Glu
                200
                                     205
                                                          210
Glu Glu Glu Ala Cys Met Ala Arg Leu Cys Val Arg Cys Ala Glu
                215
                                     220
                                                          225
Cys Leu Asp Ser Asp Tyr Glu Glu Thr
                230
<210> 42
<211> 147
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5470806CD1
<400> 42
Met Ala Ser Leu Arg Leu Phe Leu Leu Cys Leu Ala Val Leu Ile
                                                           15
Phe Ala Ser Glu Ala Gly Pro Gly Gly Ala Gly Glu Ser Lys Cys
                 20
                                      25
Pro Leu Met Val Lys Val Leu Asp Ala Val Arg Gly Ser Pro Ala
                 35
                                      40
                                                            45
Val Asp Val Ala Val Lys Val Phe Lys Lys Thr Ala Asp Gly Ser
                 50
                                      55
                                                            60
Trp Glu Pro Phe Ala Ser Gly Lys Thr Ala Glu Ser Gly Glu Leu
                 65
                                      70
                                                            75
His Gly Leu Thr Thr Asp Glu Lys Phe Thr Glu Gly Val Tyr Arg
                 80
                                      85
                                                            90
Val Glu Leu Asp Thr Lys Ser Tyr Trp Lys Ala Leu Gly Ile Ser
                 95
                                     100
                                                          105
Pro Phe His Glu Tyr Ala Glu Val Val Phe Thr Ala Asn Asp Ser
                110
                                     115
                                                          120
Gly His Arg His Tyr Thr Ile Ala Ala Leu Leu Ser Pro Tyr Ser
                125
                                     130
                                                          135
Tyr Ser Thr Thr Ala Val Val Ser Asn Pro Gln Asn
                140
                                     145
<210> 43
<211> 147
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5473242CD1
```

```
<400> 43
Met Val His Leu Thr Asp Ala Glu Lys Ala Thr Val Asn Gly Leu
   1
                                           10
 Trp Gly Lys Val Asn Pro Val Glu Ile Gly Ala Glu Ser Leu Ala
                    20
                                            25
                                                                  30
Ser Leu Leu Ile Val Tyr Pro Trp Thr Gln Arg Tyr Phe Ser Lys
                    35
                                            40
Phe Gly Asp Leu Ser Ser Val Ser Ala Ile Met Gly Asn Pro Gln
                    50
                                           55
                                                                  60
Val Lys Ala His Gly Glu Lys Val Ile Asn Ala Phe Asp Asp Gly
                    65
                                           70
                                                                  75
Leu Lys His Leu Asp
                       Asn Leu Lys Gly Thr Phe Ala Ser Leu Ser
                    80
                                           85
                                                                  90
Glu Leu His Cys Asp
                       Lys Leu His Val Asp Pro Glu Asn Phe Arg
                    95
                                          100
                                                                 105
Leu Leu Gly Asn Met Ile Val Ile Met Met Gly His His Leu Gly
                   110
                                          115
                                                                 120
Lys Glu Phe Thr Pro Ser Ala Gln Ala Phe Gln Lys Val Val
                   125
                                          130
Ala Gly Val Ala Ser Ala Leu Ala His Lys Tyr His
                   140
                                          145
<210> 44
<211> 2701
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 264114CB1
<400> 44
geggeggege cagetteete ggeeggagg gaggegagae eccaggegag geegeggeg 60 gagggeeacg ecceegaege egegeeggag gggeecagtg tggaeggge eaceggetgg 120 agegggateee acaceteegg acegagggae geggttaete eacaggatee getgaacata 180
ggatgttgcc acaaaatcta cctcgtgtat ttttctcttt cactcatgag ctgcacaatt 240
gcagatttga gcacaatgtc tgcagactgt gttgaaaaac tctgaagaac ctaattaaca 300 caggatgacc taggagtgat tctaagtctg tgtaacaaga tattactcat tagtgaatgt 360
gtcagtcttg gtactgaatg ctgcagataa cagcaagtag gttctccttt atttctgaag 420
tattcacttg accttccatc agtaagacgg acttttctaa tctgttcctg gagatattaa 480
tggaatacag tcatgtccac tcaagacgag aggcagatca atactgaata tgctgtgtca 540
ttgttggaac agttgaaact gttttatgaa cagcagttgt ttactgacat agtgttaatt 600
gttgagggca ctgaattccc ttgtcataag atggttcttg caacatgtag ctcttatttc 660
agggccatgt ttatgagtgg actaagtgaa agcaaacaaa cccatgtaca cctgaggaat 720 gtcgatgctg ccaccttaca gataataata acttatgcat acacgggtaa cttggcaatg 780
aatgacagca ctgtagaaca gctttatgaa acagcttgct tcctacaggt agaagatgtg 840
ttacaacgtt gtcgagaata tttaattaaa aaaataaatg cagagaattg tgtacgattg 900
ttgagttttg ctgatctctt cagttgtgag gaattaaaac agagtgctaa aagaatggtg 960
gagcacaagt tcactgctgt gtatcatcag gacgcgttca tgcagctgtc acatgaccta 1020
ctgatagata ttctcagtag tgacaattta aatgtagaaa aggaagaaac cgttcgagaa 1080
gctgctatgc tgtggctaga gtataacaca gaatcacgat cccagtattt gtcttctgtt 1140 cttagccaaa tcagaattga tgcactttca gaagtaacac agagagcttg gtttcaaggt 1200
ctgccaccca atgataagtc agtggtggtt caaggtctgt ataagtccat gcccaagttt 1260
ttcaaaccaa gacttgggat gactaaagag gaaatgatga ttttcattga agcatcttca 1320
gaaaateett gtagtettta etettetgte tgttacagee eecaageaga aaaagtttae 1380
aagttatgta gcccaccagc tgatttgcat aaggttggga ccgttgtaac tcctgataat 1440
gatatetaca tageaggggg teaagtteet etgaaaaaca caaaaacaaa teacagtaaa 1500
acaagcaaac ttcagactgc cttcagaact gtgaattgct tttattggtt tgatgcacag 1560 caaaatacct ggtttccaaa gaccccaatg ctttttgtcc gcataaagcc atctttggtt 1620
tgctgtgaag gctatatcta tgcaattgga ggagatagcg taggtggaga acttaatcgg 1680
aggaccgtag aaagatacga cactgagaaa gatgagtgga cgatggtaag ccctttacct 1740
tgtgcttggc aatggagtgc agcagttgtg gttcatgact gcatttatgt gatgacactg 1800
aacctcatgt actgttattt tccaaggtct gactcatggg tagaaatggc catgagacag 1860
actagtaggt cetttgette agetgeaget tttggtgata aaatttteta tattggaggg 1920
ttgcatattg ctaccaattc cggcataaga ctcccctctg gcactgtaga tgggtcttca 1980
gtaactgtgg aaatttatga tgtgaataaa aatgagtgga aaatggcagc caacatccct 2040
gctaagaggt actctgaccc ctgtgttaga gctgttgtga tctcaaattc tctatgtgtg 2100
tttatgcgag aaacccactt aaatgagcga gctaaatacg tcacctacca atatgacctg 2160
gaacttgacc ggtggtctct gcggcagcat atatctgaac gtgtactgtg ggacttgggg 2220
```

```
agagattttc gatgcactgt ggggaaactc tatccatcct gccttgaaga gtctccatgg 2280
aaaccaccaa cttatctttt ttcaacggat gggacagaag agtttgaact ggatggagaa 2340
atggttgcac taccacctgt atagtgggga agttcaggga gtgcacgcct gagttatgtg 2400
ctttgtcatt ttctttgcta aacaaaagag gctatgaaag aactaaatat gagtacataa 2460 aattctatct ttgataaatt ttatttttat gccctactta atatttgcat cagtataata 2520
tatatcagtg agtcttacag aaagatatgc ttccataata tgaaatagat tattcaataa 2580
ttgagaaact ttatgtgtaa tcatgagagt ataagaatct ggattatcta acattgttag 2640 ccctgtgtat gtacagttca aaaagttcat ttataaaagt agtttcctgt tcctagttga 2700
                                                                                     2701
<210> 45
<211> 736
<212> DNA
 <213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 1455669CB1
gagacttagc gacagacaga cgctgggacc cacgacgaca gaaggcgccg atggccgcgc 60
ctgctgagcc ctgcgcgggg cagggggtct ggaaccagac agagcctgaa cctgccgcca 120
ccagcetget gageetgtge tteetgagaa cagcaggggt etgggtacce eccatgtace 180
tetgggteet tggteceate taceteetet teatecacea ceatggeegg ggetacetee 240 ggatgteece actetteaaa gecaagatgg tagetgeeat eeetgggage etggaaceag 300
gcaatgttcg ggggaggcag gggacaggct ggaacctggt gaagtcttaa agtagactcc 360
tcctatcggg gtgtagaagg gaatctgtta atcaaacaga gcaatattag aaaggctaca 420 gaggtcaact cagtggaaca cggttctccc aaacagattt tgtaattccg aaaatccacg 480
catgogoaaa cataogoata cactocoatg ttootggaca gtttatagot accataacct 540
ggcattttcc aaaacatacc atgtagactc ttggatacac aaggtaattt tagagccaca 600 ttaggatgaa ccttttaaaa agttatgcat ttatttttat gttcccccac tggctgtatt 660
ataggacaat ttttatatgt gatatgtatt taccttagtg tgttaaataa acactggcat 720
tccaagtgtg aaaaaa
                                                                                     736
<210> 46
<211> 1826
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2084989CB1
<400> 46
ccttaggcgc cagggacagc cgagcgttac ctggtcccgg gcagcggagt tctttaccca 60
ccccagttct ggttctgacg ccctagctca ttccgcaaat ttagggcttg ggtctggctt 120 gttcccctcc ggctcgaacc acctcttctc tgagccgagc cagctaccgg ggctcctgga 180
attgccaccc ctccctgggc acccttgagg cctccgtgga gggacgtcac ggggcagagc 240
gggacgtgag cctgagtttg ctgcaggcgt gctctgtgtg gtggctgggt tctgccaatc 300
cccgtgccca ccgggtgggc gcggccggga agctcctgcc cctccctgct ggtcggcgtc 360
acgcgtgacg tcccgcgtga tggctgggag ggcccggcgg cgacagcgga ggcagagagg 420
aaggcggttc tgagagcttc agagagcgat ggaaagcaaa atgggtgaat tgcctttaga 480
catcaacatc caggaacctc gctgggacca aagtactttc ctgggcagag cccggcactt 540 tttcactgtt actgatcctc gaaatctgct gctgtccggg gcacagctgg aagcttctcg 600 gaacatcgtg cagaactaca gggccggcgt ggtgacccca gggatcaccg aggaccagct 660
gtggagggcc aagtatgtgt atgactccgc cttccatccg gacacagggg agaaggtggt 720
cctgattggc cgcatgtcag cccaggtgcc catgaacatg accatcactg gctgcatgct 780 cacattctac aggaagaccc caaccgtggt gttctggcag tgggtgaatc agtccttcaa 840
tgccattgtt aactactcca accgcagtgg tgacactccc atcactgtga ggcagctggg 900
gacageetat gtgagtgeea ceaetggage tgtggeeaeg geeetgggae teaaateeet 960
caccaagcac etgececet tggteggeag atttgtgeee tttgeagcag tggeagetge 1020
caactgcatc aacatccccc tgatgaggca gagagagctg caggtgggca tcccggtggc 1080
tgatgaggca ggtcagaggc ttggctactc ggtgactgca gccaagcagg gaatcttcca 1140
ggtggtgatt tcaagaatct gcatggcgat tcctgccatg gccatcccac cactgatcat 1200 ggacactctg gagaagaaag acttcctgaa ggtaggcgac tgtacctctc ttgtcctgga 1260
atgggcgatg gctgggagaa gtgaccaggc cccaactctc tctccagcct cgcctgattc 1320
totaagactt gocagooott otootgacco otgoaccgco tootcoacct togttoatto 1380
agcaagaatg aactgggctg gggtgaagga actctgcagg ggcaggagga gaggacaaag 1440
```

```
gaaggaaacc aacttcatca gtgttactcc agtggcttct gacacacaga aggggactgt 1500 catagtcatg cttgatctca tgctcattct tttaccccct agtgcctcca tactgagagg 1560
tacacacggg tgaacacgca cacacagaca tgaacaggac acgaaagcaa agcacaggaa 1620
caagetetgg eteatteaca gaateattta tteacaaatg tattgagtge catgeaceag 1680
gcatgtttta gggctgagga gatggcactg aacacaatgg ttatggcccc tgtcctcatg 1740
aagtttatag tetgatgeag aaaccaataa acaaggagge acceacataa atacattett 1800
agaaagtgta aaaataaaaa aaaaaa
                                                                                         1826
<210> 47
<211> 1325
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2501034CB1
<400> 47
ccacgggtcc ggttctggac tgcagttgag tggaaatggg caacggcggg cggagcggcc 60 tgcagcaggg gaaggggaac gtggatgggg tggcagcgac tcctactgct gcctcggcct 120
cctgccagta caggtgcatc gaatgcaacc aggaggccaa agagttgtac cgagactata 180
accacggtgt gctgaagata accatctgta aatcctgcca gaaacctgta gacaaatata 240
tcgagtatga tcctgttatc atcttgatta atgctatatt gtgcaaagct caggcctaca 300
gacatattet tttcaataet caaataaata tecatggaaa actetgeata ttttgtttgc 360
tttgtgaagc atacctgagg tggtggcagc ttcaagattc caaccagaat actgcccctg 420 atgacttgat cagatatgct aaggaatggg atttctatag aatgtttgcg attgctgctt 480 tagaacaaac tgcctatttt attggcattt ttaccttcct gtgggtagaa cggcccatag 500
cggcaaaaaa aaagcccaac ttcattttgc tgctgaaagc attattatta tctagctacg 600
gaaaactctt gctgattcca gctgtcattt gggaacatga ctacacatct gtgtgcctca 660 aactcattaa agtatttgtt cttacatcaa attttcaggc aattagagtg accctaaaca 720
tcaaccgtaa gctctccttc ttggccgtgt tgagtggctt actgctggaa agcatcatgg 780
tctacttctt ccagagtatg gaatgggatg ttggaagtga ttatgccatc tttaaatctc 840 aggacttctg aagagtttta ttcttcttca ctatctgtgg catgaccagc tgtatctgaa 900
agagaaaaga catgaaatat aaaccaacct cctcatttct gttgagtaaa atgaagcaaa 960
gattggaaac actttctgaa aaagaaagca atgataatag cggtggatac ccaccccac 1020
aaatgcaccc aagagacaag ccatttacat acagatattc acagtcacac atagaaacac 1080
ccacatggac acaaggaatg ttgctgcaga gactgaatga catgcaacag gtgaaggttt 1140
atacgttata cacaaggcca ggtaagcgct cataattcac acataataaa acatctaggt 1200 ttcattcctt tgacatgttt atatcttttt aatttaaatg ttgttactgg cttaaaatat 1260 tttgtgttct tacaatagaa acgcttttaa taaagtcttt cagaataaac caaaaaaaaa 1320
aaaaa
<210> 48
<211> 1832
<212> DNA
<213> Homo sapiens
<220>
<221> misc feature
<223> Incyte ID No: 2745212CB1
tgggctgtcg ttggctggag cagcggctgc gcgggtcgcg gtgctgtgag gtctgcgggc 60
gctggcaaat ccggcccagg atgtagagct ggcagtgcct gacggcgcgt ctgacgcgga 120
gttgggtggg gtagagagta ggggggggta gtcgggggtg gtgggagaag gaggaggcgg 180 cgaatcactt ataaatggcg ccgaagcagg acccgaagcc taaattccag gagggtgagc 240
gagtgctgtg ctttcatggg cctcttcttt atgaagcaaa gtgtgtaaag gttgccataa 300
aggacaaaca agtgaaatac ttcatacatt acagtggttg gaataaaaat tgggatgaat 360
gggttccgga gagcagagta ctcaaatacg tggacaccaa tttgcagaaa cagcgagaac 420 ttcaaaaagc caatcaggag cagtatgcag aggggaagat gagaggggct gcccaggaa 480
agaagacatc tggtctgcaa cagaaaaatg ttgaagtgaa aacgaaaaag aacaaacaga 540
aaacacctgg aaatggagat ggtggcagta ccagtgagac ccctcagcct cctcggaaga 600 aaagggcccg ggtagatcct actgttgaaa atgaggaaac attcatgaac agagttgaag 660 ttaaagtaaa gattcctgaa gagctaaaac cgtggcttgt tgatgactgg gacttaatta 720
ccaggcaaaa acagctettt tatetteetg ccaagaagaa tgtggattee attettgagg 780
attatgcaaa ttacaagaaa tctcgtggaa acacagataa taaggagtat gcggttaatg 840
aagttgtggc agggataaaa gaatacttca acgtaatgtt gggtacccag ctactctata 900
aatttgagag accacagtat gctgaaattc ttgcagatca tcccgatgca cccatgtccc 960
```

```
aggtgtatgg agcgccacat ctcctgagat tatttgtacg aattggagca atgttggctt 1020
atacacetet ggatgagaag ageettgett tattacteaa ttatetteae gattteetaa 1080
agtacctggc aaagaattct gcaactttgt tcagtgccag cgattatgaa gtggctcctc 1140
ctgagtacca tcggaaagct gtgtgagagg cactctcact cacttatgtt tggatctccg 1200
taaacacatt tttgttctta gtctatctct tgtacaaacg atgtgctttg aagatgttag 1260 tgtataacaa ttgatgtttg ttttctgttt gattttaaac agagaaaaaa taaaaggggg 1320
taatagetee ttttttette tttetttt tttttcattt caaaattget gecagtgttt 1380
tcaatgatgg acaacagagg gatatgctgt agagtgtttt attgcctagt tgacaaagct 1440 gcttttgaat gctggtggtt ctattccttt gacactacgc acttttataa tacatgttaa 1500
tgctatatga caaaatgctc tgattcctag tgccaaaggt tcaattcagt gtatataact 1560
gaacacactc atccatttgt gcttttgttt ttttttatgg tgcttaaagt aaagagccca 1620 tcctttgcaa gtcatccatg ttgttactta ggcattttat cttggctcaa attgttgaag 1680
aatggtggct tgtttcatgg tttttgtatt tgtgtctaat gcacgtttta acatgataga 1740
cgcaatgcat tgtgtagcta gttttctgga aaagtcaatc ttttaggaat tgtttttcag 1800
atcttcaata aatttttct ttaaatttca aa
                                                                                  1832
<210> 49
<211> 1211
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 4833111CB1
<400> 49
gacagacceg ceteaaacat ggeggegeee agegegegag gaegtgatee gettetgete 60
cggcttggat tgtagccttg acgaggtctg agcgaccatg gaccggccgg ggttcgtggc 120 agcgctggtg gctggtgggg tagcaggtgt ttctgttgac ttgatattat ttcctctgga 180
taccattaaa accaggctgc agagtcccca aggatttagt aaggctggtg gttttcatgg 240
aatatatgct ggcgttcctt ctgctgctat tggatccttt cctaatgctg ctgcattttt 300 tatcacctat gaatatgtga agtggttttt gcatgctgat tcatcttcgt atttgacacc 360
tatgaaacat atgttggctg cctctgctgg agaagtggtt gcctgcctga ttcgagttcc 420
atctgaagtg gttaagcaga gggcacaggt atctgcttct acaagaacat ttcagatttt 480 ctctaacatc ttatatgaag agggtatcca agggttgtat cgaggctata aaagcacagt 540
tttaagagag atteetttt etitggteea gitteeetta tgggagteet taaaageeet 600
ctggtcctgg aggcaggatc atgtggtgga ttcttggcag tcagcagtct gtggagcttt 660
tgcaggtgga tttgccgctg cagtcaccac ccctctagac gtggcaaaga caagaattac 720 gctggcaaag gctggctcca gcactgctga tgggaatgtg ctctctgtcc tgcatggggt 780
etggeggtea caggggetgg caggattatt tgeaggtgte tteeetegaa tggeageeat 840
cagtetggga ggtttcatet ttetggggge ttatgacega aegeaeaget tgetgttgga 900
agttggcaga aagagteett gaagcagaga caageeteae etecaettet gteaagagag 960
gggcctgcag tgcaaaccct cttccgctga gcagctgtct gaactatagg ccccagtgct 1020
gaagaccagt tgtgctaaga taccggcatg gagattgtgc catccgtggt ataggctggc 1080
tggtatgaag tcattggcct gtatgccaga gagctaagag aagaaaacgg ggtctgtggc 1140
ggtactctga acaatttcct cagaacctct taataaataa gtttggtaat gctgagaaaa 1200
aaaaaaaaa a
                                                                                  1211
<210> 50
<211> 1046
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 876677CB1
<400> 50
cccacgcgtc cgggaatgtc tttcactctc tcatactttc tcctctcccc tctcccaage 60
acatotgagt tgotgootgt tottoacact tagotocaaa cocatgaaaa attgocaagt 120
ataaaagctt ctcaagaatg agatggattc tagggtgtct tcacctgaga agcaagataa 180
agagaatttc gtgggtgtca acaataaacg gcttggtgta tgtggctgga tcctgttttc 240 cctctcttc ctgttggtga tcattacctt ccccatctcc atatggatgt gcttgaagat 300
cattaaggag tatgaacgtg ctgttgtatt ccgtctggga cgcatccaag ctgacaaagc 360
caaggggcca ggtttgatcc tggtcctgcc atgcatagat gtgtttgtca aagttgacct 420 ccgaacagtt acttgcaaca ttcctccaca agagatcctc accagagact ccgtaactac 480
tcaggtagat ggagttgtct attacagaat ctatagtgct gtctcagcag tggctaatgt 540
caacgatgtc catcaagcaa catttctgct ggctcaaacc actctgagaa atgtcttagg 600
```

```
gacacagace ttgtcccaga tettagetgg acgagaagag ategeceata geatecagae 660
tttacttgat gatgccaccg aactgtgggg gatccgggtg gcccgagtgg aaatcaaaga 720 tgttcggatt cccgtgcagt tgcagagatc catggcagcc gaggctgagg ccacccggga 780
agegagagee aaggteettg cagetgaagg agaaatgaat gettecaaat eeetgaagte 840
agcetecatg gtgetggetg agteteccat agetetecag etgegetace tgeagacett 900 gageaeggta gecacegaga agaattetac gattgtgtt cetetgeea tgaatatact 960 agaggggeatt ggtggegtea getatgataa ceacaagaag ettecaaata aageetgagg 1020
tectettgeg gtagteaget attgea
<210> 51
<211> 1660
<212> DNA
<213> Homo sapiens
<220>
<221> misc feature
<223> Incyte ID No: 2326143CB1
gctcccctgc ccaccccgcc cccgtggccg agcccgggag tcgagtggga gtcggccggc 60
cggcgcgggc agccgtgacc cccgcggggg acactgcagc cggagcccgg gaggggccgc 120
gccgccaccg tctgaactag gatgtcccga catgaaggtg tcagctgtga tgcatgttta 180 aaaggaaatt ttcgaggtcg cagatataag tgtttaattt gctacgatta cgatctttgt 240
gcatcttgtt atgaaagtgg tgcaacaaca acaaggcata caactgacca cccaatgcag 300
tgcatattaa caagggtaga ttttgattta tactatggtg gggaagcttt ctctgtagag 360
cagccacagt cttttacttg tccctattgt ggaaaaatgg gctatacgga gacatctctt 420
caagaacatg ttacttctga acatgcagaa acatcaacag aagtgatttg tccaatatgt 480
gcagcgttac ctggaggcga tcctaatcat gtcacggatg actttgcagc tcatcttaca 540
cttgaacaca gagcccctag agatttagat gaatcgagtg gtgttcgaca tgtacgtaga 600 atgtttcacc ctggccgggg attaggaggt cctcgtgctc gtagatcaaa catgcacttt 660
actageagtt ctactggtgg actttettet teteagagtt catattetee aageaatagg 720 gaagecatgg ateetatage tgagettta teteagttat caggagtgag acgttetgea 780 ggaggacage ttaatteete tggeeettee getteteagt tacaacaact geagagtgag 840
ctgcagctag aacggcagca tgcccaggca gcacggcaac aactggagac cgcacgcaac 900
gcaaccegge gtactaacac aagcagtgte accactacaa teacacaate cacageaaca 960
accaacatag ctaatacaga aagcagtcag cagactctac agaattccca gtttctttta 1020
acaaggttga atgatcctaa aatgtctgaa acggagcgcc agtccatgga aagcgagcgt 1080
gcagaccgca gcctgtttgt ccaagagctc cttctgtcca ctttagtgcg tgaagagagc 1140
tcatcctcag atgaggatga tcggggggag atggcagatt ttggtgctat gggctgtgta 1200
gatattatgc ctttagatgt tgctttagaa aacctaaatt taaaagagag taataaagga 1260
aatgageete caccacetee tetttgatga cateccaatt egeagacaat gteetetgtg 1320
ctgtatttgc caatgaaagt ggacaacaac tatcttgggt ttgtttggtg attgtaattt 1380 caggtctgtc actcttgtta cattgtgtac attcaaaagg aagagagaaa atatatatga 1440
taatcatttc cacttaacta atttttactt ctagcaggta aatgtaggta gcagtgcagg 1500
ggtgatctct gcttcctgta ccttgacatg caaaaggctc tcctaatact ccacattcaa 1560 actgaagagg aaaattgaaa tctctaatga agctgctgtg tgtatttatg aatattaatg 1620
1660
<210> 52
<211> 1110
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 2786302CB1
<400> 52
cctttattag agaaagagtg acctatctac cccttacttg gggaagactc tctgcttcaa 60
cagtttcttc tgtcttctgc tcctggacat ggcatgaaaa cggaagtttg agtatttcaa 120
ggataaagaa taactctagg acaagagttg ctttggggct aaacctatct cctgtattca 180
tttagatttg gctttaaaag cccatactat tttattagct cttctatgct ttttaaaagt 240 ttttaaaaaa attaatgtgg gcttttaat tctttctcag tgggagggtg agtttgaata 300 aacctttctt ccacatgaga agtatttac aagttgcttg tcaaatttaa aagaaaatga 300
tcaaatcttc aagaaaatga tcaaatcttc tgtgacaaaa aaatggacaa atattcacca 420
cttcagagtg ttattttcct ttttgtcata aggtgtctgg aaatgaagta tggaaatgaa 540
ataatgaata aagacccagt tttcagaatc tctccacgga gtagagaaac tcatcccaat 600
```

```
ccggaagagc ccgaagaaga agatgaagat gttcaagctg aaagagtcca agcagcaaat 660
 gcactcactg ctccaaactt ggaggaggaa ccagtcataa ctgcaagctg tttacacaag
 gaatattatg agacaaagaa aagttgcttt tcaacaagaa agaagaaaat agccatcaga 780
aatgtttcct tttgtgttaa aaaaggtgaa gttttgggat tactaggaca caatggagct 840 ggtaaaagta cttccattaa aatgataact gggtgcacaa agccaactgc aggagtggtg 900
 gtgttacaag gcagcagagc atcagtaagg caacagcatg acaacagcct caagttcttg 960
gggtactgcc ctcaggagaa ctcactgtgg cccaagctta ccatgaaaga gcacttggag 1020 ttgtatgcag ctgtggaaag actggggcaa aaaagatgct gctctcagta tttcacgatt 1080
ggtgggaggt cttaagctcc aggaacaact
                                                                                           1110
 <210> 53
 <211> 1120
 <212> DNA
 <213> Homo sapiens
 <220>
<221> misc_feature
<223> Incyte ID No: 3735780CB1
gaaatccagt tatcaaaatt gactcaagaa gagagaacct aacagaacaa taacaatgga 60
agaaattggg aacattatca caaagctatc atcctgccaa actccaggct cagatgtcac 120
aggttaaaaa aagteettea tgaaaaagaa agatettaag cagcatgatg gatteagaag 180
ctcatgaaaa gaggccacca atactaacat cttcaaaaca agatatatca cctcatatta 240
caaatgttgg tgagatgaag cattacttgt gtggctgctg tgcagccttc aacaatgtcg 300
caatcacatt teccatteag aaggteetet ttegacaaca getgtatgge atcaaaacce 360
gggatgcaat acttcagttg agaagggatg gatttcgaaa tttgtatcgt ggaatccttc 420 ccccattgat gcagaagaca actacgcttg cacttatgtt tggtctgtat gaggatttat 480
cctgccttct ccacaagcat gtcagtgctc cagagtttgc aaccagtggc gtggcggcag 540
tgcttgcagg gacaacagaa gcaattttca ctccactgga aagagttcag acattgcttc 600
aagaccacaa acatcatgac aaatttacca acacttacca ggctttcaag gcactgaaat 660
gtcatggaat tggagagtat tatcgaggct tggtgcccat tcttttccgg aatggactca 720
gcaatgtctt gtttttcggc cttcgaggtc ccattaagga gcatctgcct accgcaacga 780 ctcacagtgc tcatctggtc aatgatttta tctgtggagg tctattgggt gccatgttgg 840 gattcttgtt ttttccaatt aatgttgtaa aaactcgcat acagtctcag attggtgggg 900
aatttcagtc tttccccaag gttttccaaa aaatctggct ggaacgggac agaaactga 960 taaatcttt cagaggtgcc catctgaatt accatcggtc cctcatctct tggggcataa 1020
tcaatgcaac ttatgagttc ttgttaaagg ttatatgaaa aaaccatcag ttaagtgcca 1080
tttatcaact gaatagacct tctaagaaga aaaaaaaaa
                                                                                           1120
<210> 54
<211> 886
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 039026CB1
ggccgcggct cctgtccaga ccctgaccct ccctcccaag gctcaaccgt cccccaacaa 60 ccgccagcct tgtactgatg tcggatgcga gagcctgtgc ttaagtaaga atcaggcctt 120
attggagaca ttcaagcaaa ggttggacaa ctacttttcc agaacagaaa ggaaactcat 180 gcatcagaaa aggtgactaa taaaggtacc agaagaatat ggctgcacaa ataccagaat 240 ctgatcagat aaaacagttt aaggaatttc tggggaccta caataaactt acagagacct 300
gctttttgga ctgtgttaaa gacttcacaa caagagaagt aaaacctgaa gagaccacct 360
gttcagaaca ttgcttacag aaatatttaa aaatgacaca aagaatatcc atgagatttc 420
aggaatatca tattcagcag aatgaagccc tggcagccaa agcaggactc cttggccaac 480
cacgatagag aagteetgat ggatgaaett tigatgaaag aitgeeaaca getgetttat 540
tggaaatgag gactcatctg atagaatccc ctgaaagcag tagccaccat gttcaaccat 600 ctgtcatgac tgtttggcaa atggaaaccg ctggagaaac aaaattgcta tttaccagga 660 ataatcacaa tagaaggtct tattgttcag tgaaataata agatgcaaca tttgttgagg 720 ccttatgatt cagcagcttg gtcacttgat tagaaaaata aaccattgtt tcttcaattg 780
tgactgttaa ttttaaagca acttatgtgt tcgatcatgt atgagataga aaaattttta 840
ttactcaaag taaaataaat ggaaatatca ctgaaaaaaa aaaaaa
                                                                                           886
<210> 55
<211> 2336
```

```
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 260607CB1
taatacgctc actataggga atttggccct cgagcagtaa ttcggcacga ggaccatctc 60
tttaggatat atttttaaat tctttgaaac acataaccaa aatggtttga ttcactgact 120
gactttgaag ctgcatctgc cagttacacc ccaaatggct ttaatcccct ctcgggtctg 180
gttgcctttt gcagtttggg ttgtggactc agctcctgtg aggggtctgg ttaggagaga 240 gccattttta aggacaggga gttttatagc ccttttctac tttcctccc tcctccagt 300
ccttatcaat ctttttcct ttttcctgac ccctccttc tggaggcagt tgggagctat 360
ccttgtttat gcctcactat tggcagaaaa gaccccattt aaaacccaga gaacactgga 420 gggggatgct ctagttggtt ctgtgtccat tttcctctgt gccaaagaca gacagacaga 480 ggctgaagaga ggctgttcct gaatcaaagc aatagccagc tttcgacaca tacctggctg 540
totgaggagg aaggcotcot ggaaactggg agotaagggo gaggcootto cottoagagg 600
ctcctggggg attagggtgt ggtgtttgcc aagccaaggg gtagggagcc gagaaattgg 660 tctgtcggct cctggttgca ctttggggaa ggagaggaag tttggggctc caggtagctc 720
cctgttgtgg gactgctctg tcccctgccc ctactgcaga gatagcactg ccgagttccc 780
ttcaggcctg gcagacgggc agtgaggagg ggcctcagtt agctctcaag ggtgccttcc 840 cctcctcca acccagacat accctctgcc aaactgggaa ccagcagtgc tagtaactac 900
ctcacagagc cccagagggc ctgcttgagc cttcttgctc cacaggagaa gctggtgcct 960
ctaggcaacc ccttcctcc acctctcatc aggggtgggg gttctccttt ctttcccctg 1020 aagtgtttat ggggagatcc tagtggcttt gccattcaaa ccactcgact gtttgcctgt 1080 ttcttgaaaa ccagtagaag ggaaacagca cagcctgtca cagtaattgc aggaagattg 1140
aagaaaaatc ctcatcaatg ccaggggaca taaaagccat ttcccttcca aatactcgac 1200 aatttagatg cagaacattt ctctgtattc agacttagag taacaccagc tgaaaactgc 1260 agtttcttc ctttggatac ataaggcttc tctatcgggg tacgggacag ggaggaggcc 1320
tcatgtctga agggggattt aggggcgaga gccgcagccc tgaccctcgg tcctgtgcac 1380
cgctttgggg cacagtctga tggcgccttt gctggcgcct tagtatggtt gactccggat 1440 ggacaaaaga aaaaaaattt ttttcttga atgaaatagc aggaagctcc tcgggagcat 1500
gtgttttgat taaccgtagt gatggatgct acgagtataa atggattaac tacctcaatc 1560
cttacagtaa gattggaact aagggcaggg actcatgcat aagggtatga atcccagcca 1620 ggacaagtga gttgaggctt gtgccacaaa aggtttgtcc ttggggaaca ggcaggcctg 1680
ccaggatece ecceatateg attgggetgg gagggetgge egtgaggtee ccaettetg 1740
ctttccttgc ccatgtgtca cccctttggc ctccagcttg tccctctctc actttctata 1800 gctttgttgg accagatggt gaggaaagga atggcctctt cccttctaga gggggctggc 1860
tggagtgaga cctggggctt ggcctggaac ccaccacaca gccccaaagt caggaagcct 1920
ggggaaacca gagctgagac ctcttcaaca gggtttcttt gagatcctac acctccattg 1980
ggcccttttt cagtcttcaa tgggggccca gttggctcta gaaggagaag aggtgaagca 2040
ggatcetttg ceetggggga gtetgaggge geggteettg gaeteattea ggeegtettt 2100
gtagttgggg gagttccact gggcgatccc agcccctccc cacccacct ctaatggacc 2160 tcctcataga agccccattt cacttttgtt ttatctacct cttagcaaaa caatagataa 2220
attaggtagt ggcagctcca cttgcttagg ttaggggggg aaaaagattt cttttccaa 2280
aggaaaaaaa tattaccttg agaatacttt ccaaaaaata aaatttaaaa aaaaaa
                                                                                             2336
<210> 56
<211> 2200
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 1429651CB1
gagaacttta cgcctggatc tcatctaact gacacagaaa ccctgtaagg atccagaggt 60
ctcgttcagg accatggaga gcggcaccag cagccctcag cctccacagt tagatcccct 120
ggatgcgttt ccccagaagg gcttggagcc tggggacatc gcggtgctag ttctgtactt 180 cctctttgtc ctggctgttg gactatggtc cacagtgaag accaaaagag acacagtgaa 240
aggctacttc ctggctggag gggacatggt gtggtggcca gtgggtgcat ccttgtttgc 300
cagcaatgtt ggaagtggac atttcattgg cctggcaggg tcaggtgctg ctacgggcat 360 ttctgtatca gcttatgaac ttaatggctt gttttctgtg ctgatgttgg cctggatctt 420
cctacccatc tacattgctg gtcaggtcac cacgatgcca gaatacctac ggaagcgctt 480
cggtggcatc agaatcccca tcatcctggc tgtactctac ctatttatct acatcttcac 540
caagateteg gtagacatgt atgeaggtge catetteate cageagtett tgeacetgga 600
```

```
totgtacctg gocatagttg ggotactggc catcactgct gtatacacgg ttgctggtgg 660
cctggctgct gtgatctaca cggatgcct gcagacgctg atcatgctta taggagcgct 720 caccttgatg ggctacagtt tcgccgcggt tggtgggatg gaaggactga aggagaagta 780 cttcttggcc ctggctagca accggagtga gaacagcagc tgcgggctgc cccgggaaga 840
 tgccttccat attttccgag atccgctgac atctgatctc ccgtggccgg gggtcctatt 900
 tggaatgtcc atcccatccc tctggtactg gtgcacggat caggtgattg tccagcggac 960 tctggctgcc aagaacctgt cccatgccaa aggaggtgct ctgatggctg catacctgaa 1020
 ggtgctgccc ctcttcataa tggtgttccc tgggatggtc agccgcatcc tcttcccaga 1080
tcaagtggcc tgtgcagatc cagagatctg ccagaagatc tgcagcaacc cctcaggctg 1140 ttcggacatc gcgtatccca aactcgtgct ggaactcctg cccacagggc tccgtgggct 1200
gatgatggct gtgatggtgg cggctctcat gtcctccctc acctccatct ttaacagtgc 1260 cagcaccatc ttcaccatgg acctctggaa tcacctccgg cctcgggcat ctgagaagga 1320
gctcatgatt gtgggcaggg tgtttgtgct gctgctggtc ctggtctcca tcctctggat 1380
ccctgtggtc caggccagcc agggcggcca gctcttcatc tatatccagt ccatcagctc 1440
ctacctgcag ccgcctgtgg cggtggtctt catcatggga tgtttctgga agaggaccaa 1500
tgaaaagggt gccttctggg gcctgatctc gggcctgctc ctgggcttgg ttaggctggt 1560 cctggacttt atttacgtgc agcctcgatg cgaccagcca gatgagcgcc cggtcctggt 1620
gaagagcatt cactacctct acttctccat gatcctgtcc acggtcaccc tcatcactgt 1680
ctccaccgtg agetggttca cagagecacc etccaaggag atggtcagec acetgacetg 1740 gtttactcgt cacgaceceg tggtccagaa ggaacaagca ecaccagcag etceettgte 1800
tettacecte tetcagaacg ggatgecaga ggecageage ageageageg tecagttega 1860
gatggttcaa gaaaacacgt ctaaaaccca cagctgtgac atgaccccaa agcagtccaa 1920
agtggtgaag gccatcctgt ggctctgtgg aatacaggag aagggcaagg aagagctccc 1980 ggccagagca gaagccatca tagtttccct ggaagaaaac cccttggtga agaccctcct 2040
ggacgtcaac ctcattttct gcgtgagctg cgccatcttt atctggggct attttgctta 2100 gtgtggggtg aacccagggg tccaaactct gtttctcttc agtgctccat ttttttaatg 2160
2200
<210> 57
<211> 2823
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2069971CB1
gaaagacata cacacttcat gtaatgctac ctgcaagtct ccctagaaaa gcagtttttg 60
taggtgaaaa caatgaagcc aggtaatatt gcaaggaggc tgtaatttta gcagacctac 120
caacaacact gatgtaggaa gctcattatt ttaatttctg gagcctttta atttttctt 180
tagaaagtgt ataaataatt gcagtgctgc tttgcttcca aaactgggca gtgagttcaa 240
caacaacgac aacaacagcc gcagctcatc ctggccgtca tggagtttct tgaaagaacg 300
tatcttgtga atgataaagc tgccaagatg tatgctttca cactagaaag aaggagctgc 360
aaatgaacac ttcatagcaa tgtggaactc caacagaaac cggtgaataa agatcagtgt 420 cccagagaga gaccagagga gctggagtca ggaggcatgt accactgcca cagtggctcc 480
aagcccacag aaaagggggc gaatgagtac gcctatgcca agtggaaact ctgttctgct 540
tcagcaatat gcttcatttt catgattgca gaggtcgtgg gtgggcacat tgctgggagt 600 cttgctgttg tcacagatgc tgcccacctc ttaattgacc tgaccagttt cctgctcagt 660
ctcttctccc tgtggttgtc atcgaagcct ccctctaagc ggctgacatt tggatggcac 720
cgagcagaga teettggtge eetgetetee ateetgtgea tetgggtggt gaetggegtg 780
ctagtgtacc tggcatgtga gcgcctgctg tatcctgatt accagatcca ggcgactgtg 840
atgatcatcg tttccagctg cgcagtggcg gccaacattg tactaactgt ggttttgcac 900 cagagatgcc ttggccacaa tcacaaggaa gtacaagcca atgccagcgt cagagctgct 960
tttgtgcatg cccttggaga tctatttcag agtatcagtg tgctaattag tgcacttatt 1020 atctacttta agccagagta taaaatagcc gacccaatct gcacattcat cttttccatc 1080
ctggtcttgg ccagcaccat cactatctta aaggacttct ccatcttact catggaaggt 1140
gtgccaaaga gcctgaatta cagtggtgtg aaagagctta ttttagcagt cgacggggtg 1200 ctgtctgtgc acagcctgca catctggtct ctaacaatga atcaagtaat tctctcagct 1260
catgttgcta cagcagccag ccgggacagc caagtggttc ggagagaaat tgctaaagcc 1320
cttagcaaaa gctttacgat gcactcactc accattcaga tggaatctcc agttgaccag 1380 gaccccgact gccttttctg tgaagacccc tgtgactagc tcagtcacac cgtcagtttc 1440
ccaaatttga caggccacct tcaaacatgc tgctatgcag tttctgcatc atagaaaata 1500
aggaaccaaa ggaagaaatt catgtcatgg tgcaatgcac attttatcta tttatttagt 1560
tccattcacc atgaaggaag aggcactgag atccatcaat caattggatt atatactgat 1620 cagtagctgt gttcaattgc aggaatgtgt atatagatta ttcctgagtg gagccgaagt 1680
aacagetgtt tgtaactate ggcaatacea aatteatete eettecaata atgeatettg 1740
agaacacata ggtaaatttg aactcaggaa agtcttacta gaaatcagtg gaagggacaa 1800
```

```
atagtcacaa aattttacca aaacattaga aacaaaaaat aaggagagcc aagtcaggaa 1860
taaaagtgac tetgtatget aacgecacat tagaacttgg tteteteacc aagetgtaat 1920
gtgatttttt tttctactct gaattggaaa tatgtatgaa tatacagaga agtgcttaca 1980
actaattttt atttacttgt cacattttgg caataaatcc ctcttatttc taaattctaa 2040
cttgtttatt tcaaaacttt atataatcac tgttcaaaag gaaatatttt cacctaccag 2100
agtgcttaaa cactggcacc agccaaagaa tgtggttgta gagacccaga agtcttcaag 2160
aacagccgac aaaaacattc gagttgaccc caccaagttg ttgccacaga taatttagat 2220 atttacctgc aggaaggaat aaagcagatg caaccaattc attcagtcca cgagcatgat 2280
gtgagcactg ctttgtgcta gacattgggc ttagcattga aactataaag aggaatcaga 2340 cgcagcaagt gcttctgtgt tctggtagca actcaacact atctgtggag agtaaactga 2400
cttctgcatt tttaaaagtt acccagagat gcttctaaag atgagccata gtctagaaga 2520
ttgtcaacca caggagttca ttgagtggga cagctagata catacattgg cagctacaat 2580
agtatcatga attgcaatga tgtagtgggg tataaaagga aagcgatgga tattgccgga 2640
tgggcatggc cagtgatgtt tcacgtcatt gaggtgacag ctctgctgga ctttgaatta 2700
catatggagg ctctccagga agacgaagaa gagaaggaca ttctaggcaa aaagaagact 2760 aggcacaagg cacacttatg tttgtctgtt agcttttagt tgaaaaagca agatacaggg 2820
<210> 58
<211> 1491
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2329339CB1
<400> 58
cgcctccctc cagctgcgag tgcggcctcg gctggcggcg gcaccaggcc acagttgtaa 60
gggatcttgt ggctgtcagg atggcagagg agcaggagtt cacccagctc tgcaagttgc 120
ctgcacagcc ctcacaccca cactgcgtga acaacaccta ccgcagcgca cagcactccc 180
aggetetget eegagggetg etggetetee gggacagegg aateetette gatgttgtge 240
tggtggtgga gggcagacac atcgaggccc atcgcatcct gctggctgcg tcctgcgatt 300
acttcagagg aatgtttgct gggggattga aggagatgga acaggaagag gtcctgatcc 360 acggtgtgtc ctacaatgct atgtgccaaa tcctacattt catatacacc tccgagctgg 420
ageteageet gageaatgta caagagaeac tggtggetge etgecagett cagateccag 480
aaattateca tttetgetgt gattteetea tgteetgggt ggacgaagag aacatteteg 540
atgtctaccg gctggcagag ctgtttgact tgagccgcct gactgagcaa ctggacacct 600
atatecteaa aaactttgtg geettetete ggaetgaeaa gtaeegeeag etteeattgg 660
agaaggtcta ctcctcctc agcagcaatc gcctggaggt ctcctgcgag accgaggtat 720 atgagggggc ccttctctac cattatagcc tggagcaggt gcaggctgac cagatctcgc 780
tgcacgagcc cccaaagctc cttgagacag tgcggtttcc gctgatggaa gctgaggtcc 840
tgcagcggct gcatgacaag ctggacccca gccctttgag ggacacagtg gccagcggcc 900
tcatgtacca ccggaacgag agcctacagc ccagcctgca gagcccgcaa acggagctgc 960 ggtcggactt ccagtgcgtt gtgggcttcg ggggcattca ctccacgccg tccactgtcc 1020
tcagcgacca ggccaagtat ctaaacccct tactgggaga gtggaagcac ttcactgcct 1080
ccctggcccc ccgcatgtcc aaccagggca tcgcggtgct caacaacttc gtatacttga 1140 ttggagggga caacaatgtc caaggatttc gagcagagtc ccgatgctgg aggtatgacc 1200
cacggcacaa ccgctggttc cagatccagt ccctgcagca ggagcacgcc gacctgtccg 1260
tgtgtgttgt aggcaggtac atctacgctg tggcgggccg tgactaccac aatgacctga 1320
atgctgtgga gcgctacgac cctgccacca actcctgggc atacgtggcc ccactcaaga 1380
gggaggtgta tgcccacgca ggcgcgacgc tggaggggaa gatgtatatc acctgcggac 1440
gcaagettat teeetttagt gagggttaat tttagettge actggeegte g
                                                                             1491
<210> 59
<211> 986
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 2540219CB1
ggacgccaac ctcggtttga agtccagggc agtggctcct gcggacagcc agcataccag 60
gggccagtgc actgcattac aaccattgtg aggaatgagg gcctgtgccg tgtggctggc 120
gggcggcatg gcaggagcaa tttcttgggg gacagcgact cctatggatg tcgtgaaaag 180
```

```
tcgactccaa gctgatgggg tttatttaaa caaatataaa ggtgtcctgg actgtatctc 240
 ccagagttac cagaaggaag gtcttaaagt gtttttcaga ggcatcactg tgaacgcggt 300
gcggggcttc cccatgagtg cggccatgtt ccttgggtac gagctgtcgc tgcaggctat 360 ccgcggggac cacgcagtga cgagcccata agcgccagga ggtgaacaca ggatgactac 420 agtgttcccc tgggcctcat ctctgcatgt gaagccctga gagctgcaga tgtttgccct 480
 ttggacctcc aagtggacat caattagcaa gcgtgggcta ggatggtgca gacactgacg 540
tggcccttct gatgcctggg atgcctcatg agtcactgat tcaagccctc caaggttctg 600 atccccaatg cccactctgc taggctggca tcaaagagct ttccaagaaa tgtttggtcc 660
agctgagaag tcctgaccat gagcaccagg gagccagaaa ccacccagag aaacgttgct 720 tcactcctct gtctgaggat ggggagggg cagtgagctc tgggctcagc cactccctc 780 agctctcaagt aacacgtccc cgtgctcca gtctcctctc agcaccgacc aggttttcc 840
 ccgctcctgc acccgtggat cctgaggaca gcggtagcgc cttcctcacc gcacgctgag 900
 tecagtgegt getecteact gtgeacttat tagtgtetgt tgagtgatta aatcacatec 960
 tcaggtctgc agcaaataaa tgaaag
                                                                                                           986
<210> 60
<211> 4023
 <212> DNA
 <213> Homo sapiens
 <220>
<221> misc_feature
<223> Incyte ID No: 2722462CB1
gtccggggcg gcgcgtatgg cggcactatg ccatccacga ctcccaggcc cccagtctca 60
getetggggg tgagagttee ecetecagee eegcacacaa etgggagatg aattaccaag 120
aggcagcaat ctacctccag gaaggcgaga acaacgacaa gttcttcacc caccccaagg 180 atgccaaggc gctggcggc tacctctttg cacacaatca cctcttctac ctgatggagc 240 tggccacggc cctgctgctg ctgctgctct ccctgtgcga ggccccgcc gtccccgcac 300
teeggettgg catetatgte caegecaece tggagetgtt tgeeetgatg gtggtagtgt 360
ttgaactctg catgaagtta cgctggctgg gcctccacac cttcatccgg cacaagcgga 420 ccatggtcaa gacctcggtg ctggtggtgc agtttgtcga ggccatcgtg gtgttggtac 480
ggcagatgtc ccatgtgcgg gtgacccgag cactgcgctg cattttcctg gtggactgtc 540 ggtattgcgg tggcgtccgg cgcaacctgc ggcagatctt ccagtccctg ccgcccttca 600
tggacatcct cctgctgctg ctgttcttca tgatcatctt tgccatcctc ggtttctact 660 tgttctcccc taacccttca gacccctact tcagcaccct ggagaacagc atcgtcagtc 720
tgtttgtcct tctgaccaca gccaatttcc cagatgtgat gatgccctcc tactcccgga 780 acccctggtc ctgcgtcttc ttcatcgtgt acctctccat cgagctgtat ttcatcatga 840 acctgcttct ggctgtgtg ttcgacacct tcaatgacat tgagaaacgc aagttcaagt 900
ctttgctact gcacaagcga accgctatcc agcatgccta ccgcctgctc atcagccaga 960
ggaggcetge eggeatetee tacaggeagt ttgaaggeet catgegette tacaageece 1020 ggatgagtge cagggagege tatettacet teaaggeect gaatcagaac aacacacee 1080
tgctcagcct aaaggacttt tacgatatct acgaagttgc tgctttgaag tggaaggcca 1140
agaaaaacag agagcactgg tttgatgagc ttcccaggac ggcgctcctc atcttcaaag 1200 gtattaatat ccttgtgaag tccaaggcct tccagtattt catgtacttg gtggtggcag 1260
tcaacggggt ctggatcete gtggagacat ttatgetgaa aggtgggaac ttetteteca 1320
agcacgtgcc ctggagttac ctcgtctttc taactatcta tggggtggag ctgttcctga 1380 aggttgccgg cctgggccct gtggagtact tgtcttccgg atggaacttg tttgacttct 1440 ccgtgacagt gttcgccttc ctgggactgc tggcgctggc cctcaacatg gagcccttct 1560
atttcatcgt ggtcctgcgc cccctccagc tgctgaggtt gtttaagttg aaggagcgct 1560
accgcaacgt gctggacacc atgttcgagc tgctgccccg gatggccagc ctgggcctca 1620 ccctgctcat cttttactac tccttcgcca tcgtgggcat ggagttcttc tgcgggatcg 1680
tcttccccaa ctgctgcaac acgagtacag tggcagatgc ctaccgctgg cgcaaccaca 1740
ccgtgggcaa caggaccgtg gtggaggaag gctactatta tctcaataat tttgacaaca 1800 tcctcaacag ctttgtgacc ctgtttgagc tcacagttgt caacaactgg tacatcatca 1860
tggaaggcgt cacctctcag acctcccact ggagccgcct ctacttcatg accttttaca 1920
ttgtgaccat ggtggtgatg acgatcattg tcgcctttat cctcgaggcc ttcgtcttcc 1980 gaatgaacta cagccgcaag aaccaggact cggaagttga tggtggcatc acccttgaga 2040
aggaaatete caaagaagag etggttgeeg teetggaget etacegggag geaegggggg 2100
cctcctcgga tgtcaccagg ctgctggaga ccctctccca gatggagaga taccagcaac 2160
attocatggt gtttctggga cggcgatcaa ggaccaagag cgacctgagc ctgaagatgt 2220 accaggagga gatccaggag tggtatgagg agcatgccag ggagcaagag cagcagcgac 2280 aactcagcag cagtgcagcc cccgccgcc agcagcccc aggcagccgc cagcgctccc 2340
agaccettac ctagcccagc gcccgaaagc cgtctcttct atgcaataac acaatagtat 2400 tactctactg cgatgtacgg aactgcggtg tgtgtacaca tactcacgta tatgcacata 2400
tttatataca ggaagaaaaa agacagacaa gatggggctt ggtttataac caccttgccc 2520
tgtcttcctt aactccagaa gccagtttgg tgaggggtgg gggtgcggcc accaggtctg 2580
```

```
agetetteet aetgtggaag geteeagaag geeetteaca aggagaeeee teacetggat 2640 eeagtegaet geggggettg eeecteatgt gggetggeet ceateggeea egteeaaage 2700
tgtcactgct actgcttcag gctcacatcc ccccgacctg atggcgtgcc cgcccctct 2760
ccctgcggcc catgccacag gtttctgtgt tttgctttag ggacagaacc acttaggaaa 2820 gaaagaactc ccggtctcca gggtggtatt tcagtgtctg tgataatgtc acgcaacacc 2880
tettegggga ceagtgeeca ggatetaatg gaageggaat tggggeaact gggeeegtgt 2940
ggccagaget cagttageca gtgccgggeg gccacagatt acactgacea atetectece 3000 ttggctctgc aageetecca eccageette tetggettaa ecettgttgg egaaaaetet 3060
tccacagtgg cctccttggg gacccagaac ccggaggaag gggcatgagg caggaagtgg 3120
ggccgatgtc tgcaacccag accacttcgt ggaatgggct cttgaccaaa tcccttttt 3180
ttgcgattta cccgttcaag caaaacaacg ttttggttaa ctaaggattg tgctaaagcc 3240 gataccaggt ccttcacacg tgtgcactag gaacaggagc gaacagcaca gagagacgct 3300
ccctgtggga cgcagcagcc ccgtggcccc ggcccagttc ccagccaccc tccctggctc 3360
tgctcacacc agagatttcc atagcaggag cggttggtgc agaagtaggt tcagatgaac 3420 ctcagttaac gtcgccaccc ctcctcccac catggtaccc tgtaggagcc ctgtatgaca 3480
tctgagcgtg gtggaggtag gagggttgcc agctgcagtg accctgccac agaggcaggg 3540
tcagtgcaga ggtcgctttg gttccgcttc cctgggccac agaacggaac acagcatagg 3600 ttctgcagca ggagccgcag tggcaggatg gagggtgcga agggcaagga gtgcactgct 3660
gggcatteet ggccageece ggccetetgg tgcctgette etgtgaette agaaggcagg 3720
tggacagagc ctccctctgg ccttgtcctc ttcccagcca cagaacgggc agggtggcac 3780 ccgaccccag gggagcagta cctggtcccc cacccctcc tcccaaccac ctccaaggcc 3840
aagetgggte ceatagecag caeggeatgg tteteceett ceceettee caggteaggg 3900
gagttggaca agtagcaggt gtttgtttt aaagcacagc cctttgggaa agcaacacat 3960 tattgagact cactgtgatt cccccgggag tcagactggc tttgtctctt tctctctgga 4020
aaa
                                                                                          4023
<210> 61
<211> 2345
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2739264CB1
gaaaggagge ggaacteggt gatetgactg geggttteee ggeeggactg agaaggggag 60 egegetgege gtegeaggag taacetaett ggteteetge tttegegaca tggeetteaa 120
ttttgggget ccctcgggca cctccggtac cgctgcagcc accgcggccc ccgcgggtgg 180 gtttggagga tttgggacaa catctacaac tgcaggttct gcattcagct tttctgcccc 240
aactaacaca ggcactactg gactctttgg tggtactcag aacaaaggtt ttggatttgg 300
tactggtttt ggcacaacaa cgggaactag tactggttta ggtactggtt tgggaactgg 360
actgggattt ggaggattta atacacagca gcagcagcaa actacattag gtggtctctt 420 cagtcagcct acacaagctc ctacccagtc caaccagctg ataaatactg cgagtgtctct 480
ttctgctcca acgctgttgg gagatgagag agatgctatt ttggcaaaat ggaatcaact 540
gcaggcettt tggggaacag gaaaagggta tttcaacaat aatatteege cagtggaatt 600 cacacaagaa aateeetttt geegatttaa ggcagtaggt tatagttgea tgeecagtaa 660
taaagatgaa gatgggctag tggttttagt tttcaacaaa aaagaaacag agattcgaag 720
ccaacaacaa cagttggtag aatcattgca taaagttttg ggaggaaacc agacccttac 780 tgtaaatgta gagggcacta aaacattgcc agatgatcag acagaagttg ttatttatgt 840 tgttgaggt tcgccaaatg gtacttcaag aagagttcca gctacaacgc tatatgccca 900
ttttgaacaa gccaatataa aaacacaatt gcagcaactt ggtgtaaccc tttctatgac 960
tagaacagaa ctttctcctg cacagatcaa acagctttta cagaatcctc ctgctggtgt 1020
tgatectatt atetgggaae aggeeaaggt agataaeeet gattetgaaa agttaattee 1080
tgtaccaatg gtgggtttta aggaacttct ccgaagactg aaggttcaag atcagatgac 1140
taagcagcat caaaccagat tagatatcat atctgaagat attagtgagc tacaaaagaa 1200
tcaaactaca tctgtagcca aaattgcaca atacaagagg aaactcatgg atctttccca 1260
tagaacttta caggteetaa teaaacagga aatteaaagg aagagtggtt atgeeattea 1320
ggctgatgaa gagcagttgc gagttcagct ggatacgatt cagggtgaac taaatgcacc 1380 tactcagttc aagggccgac taaatgaatt gatgtctcaa atcaggatgc agaatcattt 1440
tggagcagtc agatctgaag aaaggtatta catagatgca gatctgttac gagaaatcaa 1500
gcagcatttg aaacaacaac aggaaggcct tagccatttg attagcatca ttaaagacga 1560
tctagaagat ataaagctgg tcgaacatgg attgaatgaa accatccaca tcagaggtgg 1620
tgtctttagt tgacagttca caaacttgtg taaaggtttg tgaaatgcat cttcttactg 1680
catcagacct teettaagaa tgaaaccgae cacatggagg gaaaaagaaa acaattettt 1740
cttggattgg ttttttgaga agtttactga caaattactg ttcatcaaat ctgaaatagt 1800
cacctcacag ctcttcaaag aaaacctttg aaagatttat atctaaaagc tgtatttact 1860
ttaaaagaag tgcataatta ccaaaattgt atgtactatt gtacattttt acaacagcat 1920
```

```
tttcttaaac ataatctgtg tttaatgatt attgtccatt gagcctgtac tctgctttcc 1980
ataccaagta aatatgaaat aatctacttt gcacataaca gaacaaacta taattacttg 2040 gctgttggag atttgtactt gagtataaat gtacaccagt ttttgtattt gtgaactcat 2100
 ctgtgggagg agtaaagaaa atccaaaagc atttaatgtt ttgtttttgt tctataaaga 2160
tatgaaaatg tattttata ttatttact tatttggaat ttacagagca cacctaagca 2220 attaggatat aacaaaacta cttaaccatt tttgcaacca ttttgtttt taagcctttt 2280
 tatttctaaa aagatgaaaa cttataaata aattcttaat ttgtaattac ttttaaaaaa 2340
aaaaa
                                                                                                           2345
 <210> 62
<211> 2085
 <212> DNA
 <213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2758310CB1
<400> 62
eggagatgtg gegacegttt etggeateat tetgagaete ggeagttget tetcaetget 60
geggeeggge etgtetgtgg gagetgeate etecteatet geaggegetg gaaaaccaga 120 caegategga catgeatgtg gttetgegge caaagcaege eetttggttg tgaactteat 180
tggccgtccc ctgctgtgtt cttcaggaga aacgtcaggg gccttcctcc aaggttctcc 300 agccccacac ccctgtggag gaaggtgctc tccaccgcgg tagtgggggc gcccctgctc 360
ctcggagccc gctatgtcat ggcagaggca cgggagaaga ggaggatgcg gctcgtggtg 420
gatggcatgg ggcgctttgg caggtctctg aaggtcggcc tgcagatctc cctggactac 480 tggtggtgca ccaatgttgt ccttcgaggg tggaagagcc caggctactt ggaggtgatg 540 tctgcgtgtc accaggggc ggctgatgcc ctggtggcag gggccatcag caacgggggc 600
ctctacgtga agctgggcca ggggctgtgc tccttcaacc acctgcttcc ccccgagtat 660
accoggacco tgcgcgtgct agaggacagg gccctcaagc ggggcttcca ggaggtggat 720 gagttgttcc ttgaggactt ccaggccctc ccccacgagc tcttccagga gtttgactac 780
cagccaattg ctgccgccag cctggcacag gtgcacagag ccaagctgca cgatggcacc 840 agcgtggctg tgaaggtgca gtacatcgac ctgcgggacc gctttgatgg ggacatccac 900 accctggage tcctgctgcg gctcgttgag gtcatgcacc ccagctttgg cttcagctgg 960
gtcctccagg acctgaaggg gaccctggcc caggagctgg acttcgagaa tgagggccgc 1020
aacgcagage getgtgegeg ggagetggeg cactteeeet acgtegtggt geeegegtg 1080 cactgggaca agtecageaa gegegtgete actgeegaet tetgegeegg etgeaaggte 1140
aacgatgtgg aggccatcag gagccagggg ctggcagtgc atgacatagc agaaaagctc 1200
atcaaggeet ttgetgagea gatatttae accggettea tecaetegga eccaeateet 1260 ggeaaegtte tggtgeggaa aggeeeggae gggaaagegg agetggtget getggaceae 1320
gggctctacc agttcctgga ggagaaggac cgcgcagccc tctgccagct gtggcgggcc 1380
atcatcctgc gggacgacgc cgccatgagg gcgcacgcag ccgcactggg ggtgcaagac 1440 tacctcctgt tcgccgagat gctcatgcag cgccccgtgc gcctggggca gctgtggggc 1500 tcgcacctac tgagccgcga agaggcggcc tacatggtgg acatggcccg cgagcgcttc 1600
gaggccgtca tggcggtgct cagggagctg ccgcggccca tgctgctggt gctgcgcaac 1620
atcaacaccg tgcgcgctat caacgtggcc ctcggcgccc ccgtggaccg ctacttcctt 1680 atggctaaaa gggctgtccg gggctggagc cgcctggcgg gcgccacgta tcggggtgtc 1740
tacggcacca gcctcctgcg ccacgccaag gtcgtctggg agatgctcaa gtttgaagtg 1800 gcgctcaggc tggagacctt ggccatgcgg ctgaccgcc tcctggctcg tgctctggtc 1860 cacctgagcc tcgtgccccc agcggaggag ctctaccagt acctggagac ctagggtgca 1920
gccgcccagg gccggcgggg ccctttcac cttgggctga cggaggtggc ggggctagag 1980 gtgtagacac cccgagccc gtgggcactc gcactgggg gctgtgacag cagctgggcc 2040 aggaggccgt gtaatgacca cacactcctc tcaagcaaaa aaaaa 2085
<210> 63
<211> 3014
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2762348CB1
<400> 63
ggaggcaaag gccaggctga ggatcagggt ggcccgggtg gcagcgggga ggcgctgcat 60
gctggaggct gtgctgagtg cccggtgcag gtgagccggt cctgcggagt tgtgccgagt 120 gcctgctgca gaccgaggct gggccaagat ggcgtctgtg tttcgaagcg aggagatgtg 180
```

```
tttgtcacaa ctgtttctcc aggtggaagc tgcatattgc tgtgtggctg agctcggaga 240
 gctcggattg gttcagttca aagatttaaa tatgaatgtg aacagctttc aaaggaaatt 300
tgtgaatgaa gtcagaaggt gtgaatcact ggagagaatc ctccgttttc tggaagacga 360 gatgcaaaat gagattgtag ttcagttgct cgagaaaagc ccactgaccc cgctcccacg 420
 ggaaatgatt accetggaga etgttetaga aaaactggaa ggagagttac aggaagecaa 480
ccagaaccag caggcettga aacaaagett cetagaactg acagaactga aataceteet 540 gaagaaaace caagaettet ttgagaegga aaccaattta getgatgatt tetttaetga 600
ggacacttct ggcctcctgg agttgaaagc agtgcctgca tatatgaccg gaaagttggg 660 gttcatagcc gggtgtgatc caacagggaa gaggatggct tcctttgagc ggttactgtg 720
gcgagtctgc cgaggaaacg tgtacttgaa gttcagtgag atggacgccc ctctggagga 780
 tectgtgacg aaagaagaaa tteagaagea catatteate atattttace aaggagagea 840
gctcaggcag aaaatcaaga agatctgtga tgggtttcga gccactgtct acccttgccc 900
agagectgeg gtggagegea gagagatgtt ggagagegte aatgtgagge tggaagattt 960
aatcaccgtc ataacacaaa cagagtctca ccgccagcgc ctgctgcagg aagccgctgc 1020
caactggcac tcctggctca tcaaggtgca gaagatgaaa gctgtctacc acatcctgaa 1080
catgtgcaac atcgacgtca cccagcagtg tgtcatcgcc gagatctggt tcccggtggc 1140 agatgccaca cgtatcaaga gggcactgga gcaaggcatg gaactaagtg gctcctccat 1200
ggccccatc atgaccacag tgcaatctaa aacagcccct cccacattta acaggaccaa 1260
taaattcaca gctggcttcc agaatattgt tgatgcctat ggtgtcggca gctaccggga 1320 gataaaccca gcccctaca ccatcatcac tttccccttc ctgttcgctg tgatgtttgg 1380 agactgtggt catggaaccg tgatgctcct ggctgcactt tggatgattc tgaatgagag 1440
acgettgete teccagaaga cagacaatga gatttggaac acettettee acgggegeta 1500 tetgateeta ettatgggea tetteteeat etacaegggt ttgatetaca atgaetgett 1560
ctccaagtcc ttgaacatct ttggctcttc ttggagtgtc caacccatgt tcagaaacgg 1620
cacatggaat actcatgtaa tggaggaaag tctatatctg cagctggacc cagccatacc 1680 aggagtgtat tttggaaatc catacccgtt tgggattgat ccgatttgga acttggcttc 1740
aaacaaactc acatttctga actcgtataa aatgaagatg tcggtgatcc tgggaattgt 1800
ccagatggtt tttggtgtca tcctcagcct tttcaatcac atatacttca gaagaactct 1860
caacatcatt ctgcaattta tccctgagat gatttttatc ctgtgtctgt ttggatacct 1920
ggttttcatg atcattttca aatggtgctg ctttgacgtc cacgtatctc agcacgcccc 1980
cagcatecte atecaettea teaacatgtt tetgtttaae tacagtgaet ettecaaege 2040
accectetae aaacateage aagaagteea aagtttettt gtggttatgg etttgattte 2100 tgtgeegtgg atgettetga ttaageegtt tattettaga geeagteate ggaaateeea 2160
getgeaggea tecaggatee aagaagatge caetgagaae attgaaggtg atageteeag 2220
cccttctagc cgttctggcc agaggacttc tgcagatacc cacggggctc tggacgacca 2280 tggagaagag ttcaactttg gagacgtctt tgtccaccaa gccatccaca ccatcgagta 2340
ctgcctgggc tgcatttcaa acacagcctc ctacctgcgg ctctgggccc tcagcctggc 2400
tcatgcacaa ctgtctgaag tgctctggac tatggtgatg aacagcggcc ttcagacgcg 2460 aggctgggga ggaatcgtcg gggttttat tatttttgcc gtatttgctg tcctgacagt 2520 agccatcctt ctgatcatgg agggcctctc tgctttcctg cacgccctgc gactgcactg 2580
ggttgagttc cagaacaagt tctatgtcgg ggatggttac aagttttctc cattctctt 2640
taaacacatc ctggatggca cagccgagga gtaggctgag ggctgcacct cccacggtgg 2700 tcaccatgcc aatgaaggaa gttcagtctt gtctttgata tcagcccctg caaggcgctc 2760
aatgggaagg ttgttcttgg ctcacctgaa gcatgaaact gtgtattatt tggacgtcag 2820 cctgtggatt tgatacgact taaccacgtc agaggaagga ctttggcaag tgatattgtc 2880
ttcatgtggg gtattaattc tcaaataata aagtaattga caaatgaggg gagaatgcta 2940
aacagatgtc ttcttgcaat attttaaata ttgtatttga gaaaataaac atctgagtca 3000
ttcaaaaaaa aaaa
                                                                                               3014
<210> 64
<211> 1726
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 3715961CB1
tgaaaatgcc tgctgcgtac caaggtatgt actagggcat ctggggtaag taaaaacaaa 60
cacatagage etgeetggag aageteatgg tetgatggaa agataageaa gaagagttaa 120 tttetaatea atatgataaa aaggteagag ageagtttet gaaaaacatg tttttgagtt 180
gagteetgaa agacaaggag atgttagtaa agcagagaag ggagaattea ttetagaaag 240
atcagacagt gtgtgggaag ggcagagtct gaaaagagca tgccccattt ggagaagcat 300 caagaagccc acgtgttaga agcaccggcc ccatgagaca aagacacagc tagagagatt 360 gactaggcca tgtcggaatg tcctcttatt ttatacatac ataagcatat agatacatat 420
agccaaagtt acctttttaa tgatcttttt tacccagtgt attctggagg tcgaatggtc 480
acatatgaac atctccgaga ggttgtgttt ggcaaaagtg aagatgagca ttatccctt 540
```

```
tggaaatcag tcattggagg gatgatggct ggtgttattg gccagttttt agccaatcca 600
actgacctag tgaaggttca gatgcaaatg gaaggaaaaa ggaaactgga aggaaaacca 660 ttgcgatttc gtggtgtaca tcatgcattt gcaaaaatct tagctgaagg aggaatacga 720
gggctttggg caggctgggt acccaatata caaagagcag cactggtgaa tatgggagat 780
ttaaccactt atgatacagt gaaacactac ttggtattga atacaccact tgaggacaat 840
atcatgactc acggtttatc aagtttatgt tctggactgg tagcttctat tctgggaaca 900
ccagccgatg tcatcaaaag cagaataatg aatcaaccac gagataaaca aggaagggga 960 cttttgtata aatcatcgac tgactgcttg attcaggctg ttcaaggtga aggattcatg 1020 agtctatata aaggctttt accatcttgg ctgagaatga ccccttggtc aatggtgttc 1080
tggcttactt atgaaaaaat cagagagatg agtggagtca gtccatttta aacccctaaa 1140
gatgcaaccc ttaaagatac agtgttcagt attattgaaa tatgggcatc tgcaacacat 1200
accccctatt atttctacct ctttaggaag acacctattc cacagagact gatttatagg 1260
gggcagcact ttatttttt ctggaaaccc aagttctctt tgactcctct ttttgtccaa 1320
aagtgatctg gtcggatctc acaaggccat ccaatgagac cccgcacagc attttctaaa 1380 gaagaatcga agcctgacca ctttcacctt gggcaagaag gtttggcctt tgagttgcta 1440
ttctatgctg aagagcctgc ttagaggagg agtaccagga gggagccagc atttcagatc 1500
tgaagtagac gataggaatg tggaagaaca catacatagt gcttaagaaa tacatttaac 1560 ctgttatgtc agtatttatc aatgaagttt gataattcac ttttctgtca ttgttaaagc 1620
gtacatactg taaattaaag ggaggtgaat ggaaattaat gaataaacat tttgagtttc 1680
cctagtgttg aaggaaggtg tactttttct tgtcagaaag ataaaa
                                                                                   1726
<210> 65
<211> 899
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5108194CB1
geggeggega egegteeggg eeggtgaggg ggegeggggg gegeeggggg ggeeeaageg 60
tcagcgcccg cgcctgtcgg gctgaactga ggaccgagtc tcctgccatt ccgagcaggc 120 ctggtatggg taatggtgtg aaggaaggcc cggtgcgatt gcatgaggat gctgaggctg 180
tectgtecte gteegtetea teaaagegtg accaeaggea agtgeteage tecetgetgt 240
ctggggcct ggctggtgcc cttgccaaaa cagcggtagc tcccctggac cgaaccaaaa 300 tcatcttcca agtgtcttca aaaagatttt ctgccaagga ggccttccgg gtcctctact 360
acacctacct caacgaggga tttctcagct tgtggcgcgg gaactcggcc accatggtgc 420
tggccgtaac cccgaaggaa atgtacagca acatctttca tgtcttcatc cgcatctcga 660
gagaagaggg gctgaagact ctctaccatg gatttatgcc caccgtgctg ggggtcattc 720 cctacgctgg cctgagcttc ttcacctatg agacgctcaa gagcttgcac agagagtaca 780
gcggccgcaa gcttattccc tttagtgagg gttaatttta gcttggcact ggccgtcgtt 840
ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcaca 899
<210> 66
<211> 643
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5503122CB1
ctttaagctg tagctgtggt ttctgcagca attttgtttt tgccttgaaa gaggtgctct 60
ggattatcac acctccatgt atgacaattt gtacctgcat ggaattgaag actcggaggc 120 tggttcagcg gattcctaca caagcaggcc gtctgactcc gatgtctctt tggaagagga 180 ccgggaagca attcgacagg agagagaaca gcaagcagct atccagcttg agagagcaaa 240
gtccaaacct gtagcatttg ccgtgaagac aaatgtgagc tactgcggcg ccctggacga 300
ggatgtgcct gttccaagca cagctatete etttgatget aaagaettte tacatattaa 360
agagaaatat aacaatgatt ggiggatagg aaggctggtg aaagagggct gtgaaattgg 420
cttcattcca agtccactca gattggagaa catacggatc cagcaagaac aaaaaagagg 480
acgttttcac ggagggaaat caagtggaaa ttcttcttca agtcttggag aaatggtatc 540
tgggacattc cgagcaactc ccacatcaac aggtgagggt tgtagttaaa ctcttttca 600
tacactgtat tccttttaaa aatatttgaa cacacatgca agc
                                                                                   643
```

```
<210> 67
<211> 2574
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5517972CB1
gcgcctccca cgccggccgc cgccgacgtg atcgctcggg cgcgccgggc gtggttgggg 60
gaaggggttg tgccgcgcga cggtctgcgt gctgtgccca ctcaaaaggt tccgggcgcg 120
caggacggaa gacgcagtgc tcgccactcc cactgagatt gagagacgcg gcaaggaggc 180
agcctgtgga ggaactgggt aggatttagg aacgcaccgt gcacatgctt ggtggtcttg 240 ttaagtggaa actgctgctt tagagtttgt ttggaaggtc cgggtgactc atcccaacat 300
ttacatcctt aattgttaaa gegetgeete egagegeaeg cateetgaga teetgageet 360
ttggttaaga ccgagctcta ttaagctgaa aagataaaaa ctctccagat gtcttccagt 420
aatgtcgaag tttttatccc agtgtcacaa ggaaacacca atggcttccc cgcgacagct 480
tccaatgacc tgaaggcatt tactgaagga gctgtgttaa gttttcataa catctgctat 540 cgagtaaaac tgaagagtgg ctttctacct tgtcgaaaac cagttgagaa agaaatatta 600 tcgaatatca atgggatcat gaaacctggt ctcaacgcca tcctgggacc cacaggtgga 660
ggcaaatett egitattaga tgiettaget geaaggaaag atecaagigg attateigga 720
gatgttctga taaatggagc accgcgacct gccaatttca aatgtaattc aggttacgtg 780 gtacaagatg atgttgtgat gggcactctg acggtgagag aaaacttaca gttctcagca 840
getettegge ttgcaacaac tatgacgaat catgaaaaaa acgaacggat taacagggtc 900
attcaagagt taggtctgga taaagtggca gactccaagg ttggaactca gtttatccgt 960
ggtgtgtctg gaggagaaag aaaaaggact agtataggaa tggagcttat cactgatcct 1020
tccatcttgt tcttggatga gcctacaact ggcttagact caagcacagc aaatgctgtc 1080
cttttgctcc tgaaaaggat gtctaagcag ggacgaacaa tcatcttctc cattcatcag 1140 cctcgatatt ccatcttcaa gttgtttgat agcctcacct tattggcctc aggaagactt 1200
atgttccacg ggcctgctca ggaggccttg ggatactttg aatcagctgg ttatcactgt 1260
gaggcctata ataaccctgc agacttcttc ttggacatca ttaatggaga ttccactgct 1320
gtggcattaa acagagaaga agactttaaa gccacagaga tcatagagcc ttccaagcag 1380 gataagccac tcatagaaaa attagcggag atttatgtca actcctcctt ctacaaagag 1440
acaaaagctg aattacatca actttccggg ggtgagaaga agaagaagat cacagtcttc 1500 aaggagatca gctacaccac ctccttctgt catcaactca gatgggtttc caagcgttca 1560 ttcaaaaact tgctgggtaa tccccaggcc tctatagctc agatcattgt cacagtcgta 1620
ctgggactgg ttataggtgc catttacttt gggctaaaaa atgattctac tggaatccag 1680 aacagagctg gggttctctt cttcctgacg accaaccagt gtttcagcag tgtttcagcc 1740 gtggaatctt ttgtggtaga gaagaagctc ttcatacatg aatacatcag cggatactac 1800
agagtgtcat cttatttcct tggaaaactg ttatctgatt tattacccat gaggatgtta 1860
ccaagtatta tatttacctg tatagtgtac ttcatgttag gattgaagcc aaaggcagat 1920 gccttcttcg ttatgatgtt tacccttatg atggtggctt attcagccag ttccatggca 1980
ctggccatag cagcaggtca gagtgtggtt tctgtagcaa cacttctcat gaccatctgt 2040 tttgtgttta tgatgattt ttcaggtctg ttggtcaatc tcacaaccat tgcatcttgg 2100 ctgtcatggc ttcagtactt cagcattcca cgatatggat ttacggcttt gcagcataat 2160
gaatttttgg gacaaaactt ctgcccaqqa ctcaatqcaa caqqaaacaa tccttgtaac 2220
tatgcaacat gtactggcga agaatatttg gtaaagcagg gcatcgatct ctcaccctgg 2280 ggcttgtgga agaatcacgt ggccttggct tgtatgattg ttattttcct cacaattgcc 2340 tacctgaaat tgttatttct taaaaaatat tcttaaattt ccccttaatt cagtatgatt 2400
tatcctcaca taaaaaagaa gcactttgat tgaagtattc aatcaagttt ttttgttgtt 2460
ttctgttccc ttgccatcac actgttgcac agcagcaatt gttttaaaga gatacatttt 2520
tagaaatcac aacaaactga attaaacatg aaagaaccca agaaaaaaaa aaaa
                                                                                               2574
<210> 68
<211> 1571
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5593114CB1
<400> 68
ggcgcggcga cagctagggt tcacggccac tggggcagag gagccgcgag aagatgtggg 60
ccggcaagcc tggaagagtt gtgactcttg ttgaagatcc tgcgggatgt gtatggggtg 240
```

```
ttgcttacag attgccagta ggaaaggaag aagaagtaaa agcatacctt gacttcagag 300
aaaaaggagg ctacagaacc acaacagtca ttttttatcc aaaagatccc acaacaaaac 360
cattcagtgt attgctatat attggaacat gtgataatcc tgattatctt ggtcctgcac 420 ctctggaaga cattgctgaa caaattttta atgcagctgg tccaagtgga agaaatacag 480
aatatetttt tgaacttgca aattetatta ggaacettgt gecagaagaa geagatgage 540
atcttttcgc tttggaaaaa ttagtaaagg aacgtttaga agggaaacag aacctcaatt 600 gcatataatt tagtcttcag agaattaact tcagtgcaca atgacaatat gatttggaaa 660
tacgtttact taaagatctt atttttaatg tagtgaggat attatttaaa cttttatttt 720
aactggaaat gtcctgaaac acatatttaa aatattggga tacagtgaaa gaaaaattca 780 aattttaata acataaagat ttcctaactt tatgttattg aacacttact cactagaagt 840
gagttettta gaaaaataca gtgaaggact cagtteagte ttgtttttat cagagtgata 900
atcatcctgt ttcacatccc aatactattt tgaaattcta aacaattaaa ccaaaattcc 960
aataaatata aggttatgcc ttcaatatat tcctatacaa ttctgtaacc atggtttaaa 1020
atacacaagc ttaaaataac atgcttagaa atacacaata atatgaacag tatttcagcc 1080
ttaattgtga atttccttgt tattcaagta ttaaatgaaa tcttttgagt ttttagccaa 1140
aaattggcat ttttaaaata cgaaaatttc cttggaatta taatgtactg tacctcttct 1200 tttttaaata aaggcatttt actatatgga aaataactca ctaaagcata aattacatta 1260
tacaaatcat gatcactaat gatgtagtct gtcattcact ttgtattaat cttataccaa 1320
aactgaaaaa gatgggctga tactacaaat taatggcaca tataatgaaa atttagtttt 1380 taaaacagct tttggaattc tttgtctgtc actatctcaa tttgtgtgtg tgtgtgtgta 1440
tatacataaa tatacatata aaattttttt tttctttgca gcctgcgtct ggccatccca 1500
caggctggaa agtgtaacct ctggcagaag ccaagaacag gcacctcctg gaattataat 1560
tttgttttgt t
                                                                                       1571
<210> 69
<211> 1549
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 044775CB1
<400> 69
cgcgcttagg caggcggtgg cgcggctgga gtgccgcggg gagggctgtg ccggttgctt 60 tctgcagccg catctcggcc agctctcctc gccgtccccg gggcgctgtg cgtctccagt 120
ecgggaccga agecgectge egtageggge ggecagatee gegteeegee teageggeeg 180
gaggacatgc gggagagaga atgagccaga gggacacgct ggtgcatctg tttgccggag 240
gatgtggtgg tacagtggga gctattctga catgtccact ggaagttgta aaaacacgac 300
tgcagtcatc ttctgtgacg ctttatatit ctgaagttca gctgaacacc atggctggag 360
ccagtgtcaa ccgagtagtg tctcccggac ctcttcattg cctaaaggtg atcttggaaa 420 aagaagggcc tcgttccttg tttagaggac taggccccaa tttagtgggg gtagcccctt 480
ccagagcaat atactttgct gcttattcaa actgcaagga aaagttgaat gatgtatttg 540
atcctgattc tacccaagta catatgattt cagctgcaat ggcaggcatt tactgtacat 600
ttctcccgag aaaagagtga gatcgtgtca tctcatgctc cccatccgca ggtcacttcc 660
tgtagaaata tggactaact taaacctcgt tttactgcaa tcacagcaac caaccccatt 720
tggcttataa agactcggtt acagcttgat gcaaggaacc gcggggaaag gcgaatgggt 780
gcttttgaat gtgttcgtaa agtgtatcag acagatggac taaaaggatt ttataggggc 840 atgtctgctt catatgctgg tatatcagag actgttatcc attttgttat ttatgaaagt 900
ataaaacaaa aactactgga atataagact gcttctacaa tggaaaatga tgaagagtct 960
gtgaaagaag catcagattt tgtgggaatg atgctagctg ctgccacctc aaaaacttgt 1020 gccacaacta tagcatatcc acatgaagtt gtaagaacaa gactacgtga agagggaaca 1080 aaatacagat cttttttca gactctatct ttgcttgttc aagaagaagg ttatgggtct 1140
ctttatcgtg gtctgacaac tcatctagtg agacagattc caaacacagc cattatgatg 1200 gccacctatg aattggtggt ttacctactc aatggatagc agcacgagga ctgctgtact 1260
gcaaaaaag aagaccaaaa gattacagtg gaccatggga tacagaagcc agcatggcag 1320
acagaagaaa aatagtttgg gaacatgtaa ctattctaag tggaagtttt gttgtaggaa 1380 ttatagtaat cacaccacat tacttggcct ttcggtaatg tgaaaaaaaa aaaaaaacct 1440
cagageetee aaggaaatge etttagaage acteetete caaaattgee atttteteta 1500
ccatgtcccc cagacacagt tgggttttgt tgatttatgg cagtcttct
                                                                                       1549
<210> 70
<211> 2237
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
```

```
<223> Incyte ID No: 116588CB1
<400> 70
gtcactggtg ccacggggcc tcagcgcact tctgtcttaa gctcctgggc ctccttattt 60
tcccctttgc gatcgattcc agccacacct gtggatgttg ctagttactc cgcgtccgga 120 acgtggaggt cgagggactg agctgggcga gttttgtggc actcctttgc tcttcagcag 180
ctatttttgc tatgataatc ctgctgccct tcagactcaa gttaaacgag atatgcaagt 240
gaataccacg aaattcatgc tgctgtatgc ctggtattct tggcccaatg tagttttgtg 300 tttctttggt ggcttttga tagaccgagt atttggaata cgatggggca caatcatttt 360
tagetgettt gtttgeattg gaeaggttgt ttttgeeetg ggtggaatat ttaatgettt 420
ttggctgatg gaatttggaa gatttgtatt tgggattggt ggcgagtcct tagcagttgc 480 ccagaataca tatgctgtga gctggtttaa aggcaaagaa ttaaacctgg tgtttggact 540
tcaacttagc atggctagaa ttggaagtac agtaaacatg aacctcatgg gatggctgta 600
ttctaagatt gaagetttgt taggttetge tggteacaca acceteggga teacacttat 660
gattgggggt gtaacgtgta ttctttcact aatctgtgcc ttggctcttg cctacttgga 720
tcagagagca gagagaatcc ttcataaaga acaaggaaaa acaggtgaag ttattaaatt 780
aactgatgta aaggacttct ccttacccct gtggcttata tttatcatct gtgtctgcta 840 ttatgttgct gtgttccctt ttattggact tgggaaagtt ttctttacag agaaatttgg 900
attttcttcc caggcagcaa gtgcaattaa cagtgttgta tatgtcatat cagctcccat 960
gtccccggtg tttgggctcc tggtggataa aacagggaag aacatcatct gggttctttg 1020 cgcagtagca gccactcttg tgtcccacat gatgctggcc tttacgatgt ggaacccttg 1080 gattgctatg tgtcttctgg gactctccta ctcattgctt gcctgtgcat tgtggccaat 1140
ggtggcattt gtagttcctg aacatcagct gggaactgca tatggcttca tgcagtccat 1200
tcagaatctt gggttggcca tcatttccat cattgctggt atgatactgg attctcgggg 1260 gtatttgttt ttggaagtgt tcttcattgc ctgtgtttct ttgtcacttt tatctgtggt 1320
cttactctat ttggtgaatc gtgcccaggg tgggaaccta aattattctg caagacaaag 1380
ggaagaata aaattttccc atactgaatg agaagttaaa atgaatgtgt catgagaatg 1440 ggcttaacac atcgttggtt tgaaaacttc cattttaaa aatttagagt ttagtcatta 1500
gaaaaaataa tggactggaa agttatattt atatccaaat atacctattt caaagtgtat 1560
ttgtgaggcc tgttttagcc tgtgtctttt gtattgtgtg ttgctaaaga attctacttt 1620 tagtaggcta atcaacaatg aaagggttag aaaattgctg tggaacatcc aggtgaactt 1680 caggaaagac agtgaaaaat ggaaaacgtt ggagcttctg ttgagataat cttcattagg 1740
tatatatett agggatacag cettteett atettatage aggaaaaaaa aacttetgag 1800 ggaaatagaa gggetgegtt acacaaaata aacaatggca ttgtcatagg cetteettt 1860
actagtaggg cataatgcta gggaatatgt gaagatgttt ttatgaagtc tctttctgat 1920
cacgaacaat agettgeget etactetgta gttatgtgga ttgeegagea atgaccettt 1980
tcaatttett atttetgtgt tactgaggac cetaateaet tagggatgta attttatagt 2040
ataaactttc tgtacagttt ttcttatagt ctaataagta aaaagtgtcc ttcaaattat 2100
gataattgcc tatgtacatg gataaattaa aacactgcac acggagtaaa aaaaaaaaa 2160 aaaaaaaaaa aaaaaatgag cggccgcaag cttattccct ttagtgaggg ttaattttag 2220
cttggcactg gccgtcg
                                                                                              2237
<210> 71
<211> 1114
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 875369CB1
<400> 71
tgggaacgga gcagccccgg gggccccctt gaggcggcga ggccgcgaag ggcgcggggc 60
tggagggtag gagagcgcg gaaagcgccc cagacgccac tcgcggcgga cggcggccag 120
ttcccagggg tttggagccg gcccgcggcg ccatggctca cgtcggctcc cgcaagcgct 180
cgaggagtcg cagccggtcc cggggacggg ggtcggaaaa gagaaagaag aagagcagga 240
aagacacete gaggaactge teggeeteea cateecaagg tegeaaggee ageaeggeee 300
ctggggcgga ggcctcacct tctccctgca tcacagagag aagcaagcag aaggcccgga 360
ggagaacaag atccagctcc tcctcctct cttccagttc ttctagctcc tcttcttcct 420 cctcgtcctc ctcctcttcc tccagtgatg gccggaagaa gcgggggaag tacaaggaca 480
agaggaggaa gaagaagaag aagaggaaga agctgaagaa gaagggcaag gagaaggcgg 540
aagcacagca ggtggaggct ctgccgggcc cctcgctgga ccagtggcac cgatcagctg 600
gggaggaaga ggatggccca gtcctgacgg atgagcagaa gtcccgaatc caggccatga 660 agcccatgac caaggaggag tgggatgccc ggcagagcat catccgcaag gtggtggacc 720
ctgagacggg gcgcaccagg cttattaagg gaqatggcga ggtcctagag gaaatcgtaa 780
ccaaagaacg acacagagag atcaacaagc aagccacccg aggggactgc ctggccttcc 840 agatgcgagc tgggttgctt ccctgagggc ccccgctggc caaggcctgt ggacgacgct 900
ggcggcccag cctgggcagg tttcagggtg ccagtgggaa gcctgatggg tgctggtggc 960
```

```
ctttcccccg tggattggtc tctggcccag cccagtctct tctcaggggc agggggtgga 1020
ggttggggtc accggcctgc ttggcaccc catctgaaag agcagcactt ctcagctatt 1080
aaaggccccc tggatagaca aaaaaaaaaa aaaa
<210> 72
<211> 998
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 1325518CB1
<400> 72
gactaaaaaa gccatgtatt ctttcgtttc tctctaaaag aagaaaaata taatttaaaa 60
atacattgcg tatttictaa aacaataaat ttatagtgti aatattcata gggtcaatca 120
aaatgaagct totootttgg gootgoattg tatgtgttgc ttttgcaagg aagagacggt 180
tccccttcat tggtgaggat gacaatgacg atggtcaccc acttcatcca tctctgaata 240 ttccttatgg catacggaat ttaccacctc ctctttatta tcgcccagtg aatacagtcc 300
ccagttaccc tgggaatact tacactgaca cagggttacc ttcgtatccc tggattctaa 360
cttctcctgg attcccctat gtctatcaca tccgtggttt tcccttagct actcagttga 420 atgttcctcc tctccctcct aggggtttcc cgtttgtccc tccttcaagg ttttttcag 480
cagetgeage accegetgee ceacetattg cagetgagee tgetgeaget geacetetta 540
cagecacace tgtageaget gagectgetg caggggeece tgttgeaget gagectgetg 600 cagaggeace tgttggaget gagectgetg cagaggeace tgttgeaget gagectgetg 660
cagaggcace tgttggagtg gagccagetg cagaggaace ttcaccaget gagcctgcta 720
cagccaagcc tgctgcccca gaacctcacc cttctccctc tcttgaacag gcaaatcagt 780
gaaattetet agaagagtae catgggttea tttetataet gatgeagaaa taagtgaaat 840
ctacaaaagt tttctttctt ttccaaagac tatttcattc tgttgtattc agagtattca 900
tctcactaca ttgatttgtt tgtggtagtt tttccttgga cttaatttat attgaaaaaa 960
cattgataat taaataaata aaatagataa tttagaca
                                                                                    998
<210> 73
<211> 2348
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 2060987CB1
<400> 73
gggccagccg gctcgcccgg gggccatggc agcagcggct actgcagccg agggggtccc 60
cagtcggggg cctcccgggg aagtcatcca tctgaatgtg ggaggcaaga gattcagtac 120 ctctcgccag actctcacct ggatcccaga ctccttcttc tccagtcttc tgagcggacg 180
catctcgacg ctgaaagatg agaccggagc aatcttcatc gacagggacc ctacagtctt 240 cgcccccatc ctcaacttcc tgcgcaccaa agagttggat cccaggggtg tccacggttc 300
cagcetecte catgaagee agttetatgg geteacteet etggttegte geetgeaget 360
tcgagaggag ttggatcgat cttcttgtgg aaacgtcctc ttcaatggtt acctgccgcc 420
accagtitte ceastgaage ggeggaaceg geacageeta gtggggeete ageagetagg 480 aggaeggeea geecetitee gaeggageaa caegatigeee ceeaacettig geaatgeagg 540
gctgctgggc cgaatgctgg atgagaaaac ccctccctca ccctcaggac aacctgagga 600
gccggggatg gtgcgcctgg tgtgtggaca ccataattgg atcgctgtgg cctataccca 660 gtttctagtc tgctacaggt tgaaggaagc ctctggctgg cagctggtgt tttccagccc 720
ccgcctggac tggcccatcg aacgactggc gctcacagcc cgggtgcatg gtggggcttt 780
gggtgaacat gacaagatgg tggcagcagc caccggcagc gagatcctgc tatgggctct 840 gcaggcggaa ggcggtggct ccgagatagg ggtctttcat ctgggggtgc ctgtggaggc 900 cttgttcttc gtcgggaacc agctcattgc tacaagccac acagggcgca tcggggtgtg 960
gaatgccgtc accaagcact ggcaggtcca ggaggtgcag cccatcacca gttatgacgc 1020
ggcaggetee tteeteetee tgggetgcaa caacggetee atttactacg tggatgtgca 1080
gaagtteece ttgegeatga aagacaacga ecteettgte agegagetet ategggacee 1140
aggggaggat ggggtcaccg ccctcagtgt ctacctcacc cccaagacca gtgacagtgg 1200
gaactggatc gagatcgcct atggcaccag ctcaggggc gtgcgggtca tcgtgcagca 1260 cccggagact gtgggctcgg ggcctcagct cttccagacc ttcactgtgc accgcagccc 1320
tgtcaccaag atcatgctgt cggagaagca cctcatctca gtctgtgccg acaacaacca 1380
cgtgcggaca tggtctgtga ctcgcttccg cggcatgatt tccacccagc ccggctccac 1440
cccacteget teetttaaga teetggetet ggagteggea gatgggeatg geggetgeag 1500
tgctggcaat gacattggcc cctacggtga gcgggacgac cagcaagtgt tcatccagaa 1560
```

```
ggtggtgccc agtgccagcc agctcttcgt gcgtctctca tctactgggc agcgggtgtg 1620
ctccgtgcgc tccgtggacg gctcacccac gacggccttc acagtgctgg agtgcgaggg 1680
ctcccggcgg ctcggctctc ggccccggcg ctacctgctc actggccagg ccaacggcag 1740 cttggccatg tgggacctaa ccaccgcat ggacggctc ggccaggccc ctgcaggtgg 1800
cctgacggag caagagctga tggaacagct qqaacactqt qaqctqqccc cgccggctcc 1860
ttcagctccc tcatggggct gtctccccag ccctcaccc cgcatctccc tcaccagcct 1920
ccactcagcc tccagcaaca cctccttgtc tggccaccgt gggagcccaa gcccccgca 1980
ggctgaggcc cggcgccgtg gtgggggcag ctttgtggaa cgctgccagg aactggtgcg 2040
gagtgggcca gacctccgac ggccacccac accagccccg tggccctcca gcggtctcgg 2100 cactccctc acacctccca agatgaagct caatgaaact tccttttgaa caacgcagct 2160
gccatgatgc cttgggatgc cctggtcctg ggggactcag gtgcctccct gattcctgtg 2220
ggaacccegg gttcagggcc agggcctcct tggaataaat ggttattgtt actaggtccc 2280
caccttccct cttttctgga agccaaagtc accttcccca ataaagtcct cactgccaaa 2340
aaaaaaaa
<210> 74
<211> 1139
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2172064CB1
<400> 74
ctcgagctgg gatgtgtggc aggttcctgc ggcggctgct ggcggaggag agccggcgct 60
ccaccccgt ggggcgcctc ttgcttcccg tgctcctggg attccgcctt gtgctgctgg 120
ctgccagtgg gcctggagtc tatggtgatg agcagagtga attcgtgtgt cacacccagc 180
agcoggetg caaggotgoo tgcttcgatg cottocacc cotococcg etgcgttcot 240
gggtcttcca ggtcatcttg gtggctgtac ccagcgccct ctatatgggt ttcactctgt 300
atcacgtgat ctggcactgg gaattatcag gaaaggggaa ggaggaggag accctgatcc 360
agggacggga gggcaacaca gatgtcccag gggctggaag cctcaggctg ctctgggctt 420 atgtggctca gctgggggct cggcttgtcc tggaggggc agccctgggg ttgcagtacc 480
acctgtatgg gttccagatg cccagctect ttgcatgtcg ccgagaacct tgccttggta 540
gtataacctg caatctgtcc cgccctctg agaagaccat tttcctaaag accatgtttg 600 gagtcagcgg tttctgtctc ttgtttactt ttttggagct tgtgcttctg ggtttgggga 660
gatggtggag gacctggaag cacaaatctt cctcttctaa atacttccta acttcagaga 720
gcaccagaag acacaagaaa gcaaccgata gcctcccagt ggtggaaacc aaagagcaat 780
ttcaagaagc agttccagga agaagcttag cccaggaaaa acaaagacca gttggaccca 840 gagatgcctg agttggagat gaactttggc caactttcct catcaccata cttaaaatcc 900
tgtccaagga ggagcttatt caccattttc tatacatgtg acatatgtag cagcataacc 960
gacaactggg actgcgctgc cttgactcca cctctacata caatgactca gctaaccaga 1020
ctaataaaag ccatgtttgc accattgctc agggaggcat tgctttgggg aattattccc 1080
agtgtcctcc ttacttatcg caagtaataa aaatcccctg ggaaatcctc aaaaaaaaa 1139
<210> 75
<211> 863
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2219267CB1
<400> 75
teagagetag etteegeege ageetggeac eteaggtgee gaeettetae etgeeetaet 60
tectggagge eggeetggag geggeageet tggtetteet geteetgaeg geagattget 120
gtggacgccg ccccgtgctg ctgctgggca ccatggtcac aggcctggca tccctgctgc 180
tectegetgg ggeceagtat etgecagget ggaetgtget gtteetetet gteetgggge 240
tectggeete eegggetgtg teegeactea geageetett egeggeegag gtetteecea 300
cggtgatcag gggggccggg ctgggcctgg tgctggggc cgggttcctg ggccaggcag 360 ccggcccct ggacacctg cacggccggc agggcttctt cctgcaacaa gtcgtcttcg 420
cetecettge tgteettgee etgetgtgt teetgetget geetgagage egaageeggg 480
ggctgcccca gtcactgcag gacgccgacc gcctgcgccg ctccccactc ctgcggggcc 540 gcccccgcca ggaccacctg cctctgctgc cgccctccaa ctcctactgg gccggccaca 600
cccccgagca gcactagtcc tgcctggtgg ccctgggagc caggatggga ccaaagtcaa 660
ggcctggggc atggctgagt accccagacg tctggtccag ggcagacaca ttcctctcag 720
aagcccgtgt ctcagtgcag gtggagccgt ggggacagcg tgaaggtgtc tccagccagg 780
```

٠. <u>۽</u>

```
ccccaggcac tgggaggccc tgggtctccc cccagccaca cccagtaggt gtggaggata 840
aaggettetg tggaaaaaaa aaa
<210> 76
<211> 1322
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 2308629CB1
<400> 76
ccggggggcc ggcgcgggg gaggccgggg cctgcaggcc cccggtacga caagatccgg 60
actccggccc ggactacgag gcgctgccgg ctggagccac tgtcaccacg cacatggtgg 120 caggcgccgt ggcagggatc ctggagcact gcgtgatgta ccccatcgac tgcgtcaaga 180
cccggatgca gagtctacag cctgacccag ctgcccgcta tcgcaatgtg ttggaggccc 240
tetggaggat tataagaacg gagggeetat ggaggeecat gagggggetg aacgteacag 300 caacaggege agggeetgee caegeeettt attttgeetg etacgaaaag ttaaaaaaga 360
cattgagtga tgtaatccac cctgggggca atagccatat tgccaatggt gcggccgggt 420
gtgtggcaac attacttcat gatgcagcca tgaaccctgc ggaagtggtc aagcagagga 480 tgcagatgta caactcacca taccaccggg tgacagactg tgtacgggca gtgtggcaaa 540
atgaaggggc cggggccttt taccgcagct acaccaccca gctgaccatg aacgttcctt 600
tccaagccat tcacttcatg acctatgaat tcctgcagga gcactttaac ccccagagac 660
ggtacaaccc aagcteccac gteetetetg gagettgege aggagetgta getgeegeag 720 ccacaaccc actggacgtt tgcaaaacac tgetcaacac ecaggagtec ttggetttga 780
actcacacat tacaggacat atcacaggca tggctagtgc cttcaggacg gtatatcaag 840
taggtggggt gaccgcctat ttccgagggg tgcaggccag agtaatttac cagatcccct 900 ccacagccat cgcatggtct gtgtatgagt tcttcaaata cctaatcact aaaaggcaag 960
aagagtggag ggctggcaag tgaagtagca ctgaacgaag ccaggggttc agatgacact 1020 gctgcatcct ggtcacattc tctgtctcct ggaatgctcc cacctcaagt ggagttagaa 1080 ggaaggtaga ggggctctcc cccaggattt tggtgttttg actaacacca gttcctgcca 1140
acctctgttg ccaccacctt tccttccagg ccctaagcac gtgcagcaaa gcacaccaca 1200
aa
                                                                                           1322
<210> 77
<211> 1869
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2660038CB1
cagggacgct cattcctggc cagagtcctg acttctttct cggatcgaga tctcgttgct 60
ggctcgagga aatcaccggc tcttctcccg gatccttttc ttctctcact tgctggcttt 120
cttcttcctg cctccctggc ttccatctcc caaccccgcg attccctcct ctactcccgt 180 gctcccgtcg cccgccatcc tgagccatcc cacctgcaac cttctgtctt ttgcccctcc 240
ttgacctcag agggtcctgc cttaagcttc tcaccagaat ctcctagatt tctatctctt 300
ccctgcttgc cagctcttga ctcccaaatt ccagctgacg tttgaccact tgacatttga 360
ccctgacacc cttgactgca aatctaaatt cttatcttct gcaacctgta ctgctgaaac 420 aggccttccc cctgtcttcc aaccccgctt tctgacaccc atttttactt tcttactctt 480
gggtcagttc ctgctacagc tatcccacca tggacttctt gatgagtggc ctggcagcct 540
geggggeetg tgtatteace aateceetgg aggtggtgaa gaceaggatg cagttgeaag 600 gagaaetgea ggeeeetgge acataceage ggeaetaceg aaatgtette catgeettea 660
tcaccatcgg caaggtggat ggccttgctg ccctgcagaa aggcctggcc cccgcctct 720
tgtaccagtt cctgatgaat ggcatccgac tgggcaccta tgggctggct gaggctgggg 780 gctacctgca cacagccgaa gcgacccaca gtcctgcccg cagcgcagca gctggggcca 840
tggctggggt catgggagcc tacttgggga gccccatcta catggtgaag acacacctgc 900
aggcacaggc agcctcagaa attgctgtag ggcaccagta taagcatcag ggcatgtttc 960
aggcgctaac cgagattggc cagaaacatg gtctggtggg gttatggcgt ggggctctgg 1020 gcggcctgcc ccgagttatc gtcggttcct ccacccagct gtgcaccttc tcatccacca 1080
aggacetect gagecagtgg gagatettte etceecagag etggaagttg gegetggtgg 1140 etgecatgat gagtggeatt geagttgtet tggccatgge accetttgat gtggeetgea 1200
caaggeteta caaccageee acagatgeae agggeaaggg ceteatgtae egggggatae 1260
```

```
tggacgctct gctgcagaca gctcggaccg agggcatttt tggcatgtac aagggtatag 1320
gtgcctccta cttccgcctc ggcccccaca ccatcctctc cctcttcttc tgggaccagc 1380
tgcgctccct ctactacaca gacactaaat aacagccgct ttcccagtct ccaccaaatg 1440
agractectt ggccacttgt gcctccacca ctatgtcctg gtgactactg attaggtgac 1500 ctttcatcca tccatggggg acagccaacc ccactccca tctgttctca gggttgaatc 1560
actacaagag atgagtttcc cttctttcct tgggtgttgc tttaaacctt ccctacccat 1620
tccctgggta actcacacc ctctctcagg gctgaacgag tcatcccaaa gtgtatttcc 1680 tcccactcac cactgccacc cttgagtccc tcctgctccc atgcacagtt ttaaactcct 1740
ccctccaaaa ccaaagggaa tcgagagacc caattcccag gcgtctggga cccaggtgtc 1800
ctgttagatt caaaggcaca gagattatat tgattataaa gcaagtttat tctgaaaaaa 1860
aaaaaaaa
                                                                                        1869
<210> 78
<211> 1881
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2670745CB1
gaagaaccga gcttggctgt gtttatctcg ttggggacta aggcgtcggt tggcgcgcaa 60 cgggttctag gctgcaggca gctcgaggac ccgcggccc gcccggctc ggcctggcag 120
atagcagagg cagcaggccg tgccgggggg gcatgttgct gtaaccagtg gcccagggga 180
tgttacggtg gacagtgcac ctggagggcg ggccccgcag ggtgaaccat gctgcagtgg 240 ctgtcgggca tcgggtatac tccttcgggg gttactgctc tggtgaagac tatgagacac 300
tgcgtcagat agatgtgcac attttcaatg cagtgtcctt gcgttggaca aagctgcccc 360
cggtgaagtc tgccatccgt gggcaagctc ctgtggtacc ctacatgcgc tatggacact 420 caaccgtcct catcgacgac acagtcctcc tttggggcgg gcggaatgac accgaagggg 480
cctgcaatgt gctctatgcc tttgacgtca atacgcacaa gtggttcaca ccccgagtgt 540
cagggacagt tectggggee egggatggae atteageetg tgteetagge aagateatgt 600
acatttttgg gggctacgag cagcaggcgg actgtttttc caatgacatt cacaagctag 660
ataccagcac catgacatgg acticttatet gtacaaaggg cagccetgca cgetggaggg 720
acttccactc agccacaatg ctgggaagtc acatgtatgt ctttgggggc cgtgccgacc 780
getttgggee attecattee aacaatgaga tttactgeaa eegeattega gtetttgaca 840 eeagaactga ggettggetg gaetgteee egaeteeagt getgeetgag gggegeegga 900
gccactcggc ctttggctac aatggggagc tgtacatctt tggtggttat aatgcaaggc 960
tgaaccggca cttccatgac ctctggaagt ttaatcctgt gtcctttacc tggaaaaaga 1020 ttgaaccgaa ggggaagggg ccatgtccc gccggcgcca gtgctgctgt attgttggtg 1080
acaagattgt cctctttggg ggtaccagtc catctcctga ggaaggcctg ggagatgaat 1140 ttgaccttat agatcattct gacttacaca ttttggactt tagccctagt ctgaagactc 1200 tgtgcaaact ggccgtgatt cagtataacc tagaccagtc ctgtttgcct catgatatca 1260
ggtgggagct gaatgccatg accaccaaca gcaatatcag tcgccccatc gtctcctccc 1320
atgggtagga ggaagtttct gccacctccc ctcctgagcc tgctgtcatc ttcactgccc 1380 ctgcccatct gtcacccacc tgctcctttg acccctggac ttggtatacc tccatgtgga 1440
gttgttgggc gagaggtgtt ctctgtgctg tgaattcagt ggggagctgt agcggggtgg 1500
gggctaggtt cctccccct tgggccgagg gcccttccc cttggtgctc tgtccccatc 1560
cacctccttt cagctgctcc tgggcctcag ctctgcccag ggccagccag gttctgctgg 1620 gaagggaagg gaatggggag aagggagaag caagcagtgt ctgagcctca ggagcttccc 1680
ceteceett tgeetateee eteceetetg ettgageett gageettgae tgggagetga 1740
aaggagttgc agctgttggc atgagacctc cttctccccg tcttggggag gtggggacca 1800 gcagataaat cccaccttc cttgagctgt cgctgtactc tgaagttcag ccagctcaga 1860
ttttataaaa attaattaaa a
                                                                                         1881
<210> 79
<211> 2004
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 2676443CB1
<400> 79
aatcgcaggg ggcgtggcct agcggtcccg cccccggggc gcgcgcgcc attggctgtg 60
cggggtgcgg gcgcgcgggc ggcgctttga accgggcgcg gggcgcgggg cgcggggcgc 120
tgcggccggt acacgccggg gtagggccgg ggtcgggttg tggtcgggcc gggattgggc 180
```

```
tetectggge catggeagee gaggegege tgtegegetg gtaetteggg gggetggeet 240
cetgegggge egeetgetge acgeaceege tggacetget caaggtgeat etgeagaege 300 ageaggaggt gaagetgege atgaegggea tggegetgeg ggtggtgegt accgaeggea 360
teetggeact ctacagegge etgagegeet egetgtgeag acagatgace tacteeetga 420
ctcggttcgc catctacgag actgtgcggg accgcgtggc caagggcagc caggggcctc 480 tccccttcca cgagaaggtg ttgctgggct ccgtcagcgg tttagctgga ggcttcgtgg 540
ggacgcccgc agacttggtc aacgtcagga tgcagaacga cgtgaagctg ccccagggtc 600
ageggegeaa ctaegeeeat gegetggatg geetgtaeeg egtagetegt gaagagggte 660 teaggagaet gttetegggt geaaceatgg cateeageeg aggggeetta gteaetgtgg 720
gecagetgte etgetacgae caggecaage agetggteet tageaceggg tacetetetg 780
acaacatett cacteaettt gtegeeaget ttattgeagg tggatgtgee aegtteetgt 840
gccagcccct ggatgtgctg aagactcgcc tgatgaactc caagggggag tatcagggcg 900
ttttccactg cgccgtggag acagcgaaqc tcggqcctct gqccttttac aagggctcg 960
teccagetgg cateegeete atececeaca eegtgeteac tittgtgttt etggaacage 1020
tacgcaaaaa ctttggcatc aaagtgccat cctgaccagc cgtgggaatg gctgggctgc 1080 caggccagac acgctaggtt cttccaaaga gtcccaagcc cagcacctgc tcctggggcc 1140
acgaeeteee tggeegtgge cacceateet eegeageagg eeeetgetgt eeeeecaeet 1200
gctggctgag ctcccctgg cctcgtcccc tctcagctgt agctgcacca cccccgctct 1260 ggctaccagg ctctcccggc tgggcactgc gtggccttgc ccctctcccg ctggcagctc 1320
tattatccct gcctcctgcc cccgatgccc aaagcagcat cttccagcac tttccatcga 1440 ggacttgggt ggcagagtgt gggtgcagcc tggctgttgc tcacccaagt gctagctctg 1500
cacttegtgt etgetgagag caaccagace ttecatgtee tegggeaget geaacteece 1560
gcgagacccc gcagctgggt gggatgaaca agcaacgcag accacaagcg agtgcctggg 1620 agggagtggc ccagggtggt tctggagcca ttgtgggtga gggtcgaggg ccaccgaggt 1680
cccgcgcacc getgcctgcc ctgcagtggc tttaacagtt agttttgcca aagcctctcc 1740
actcaccage aggeggtete tgtetteagg gattgtgeet gegteeeteg ggeacetggg 1800 cececeget tggeteeetg ggggaatgge ecaggeggge egeggteet eettagggee 1860
ttctccccga caaggagtcc gacggggcgg atgctgcatc ctctgcctcc ctggtcgctg 1920
ggetteacce cacetgggaa gggcagtgtg etetgtgggg getgeaatea ataaatgeeg 1980
qqaqctqcca aaaaaaaaaa aaaa
                                                                                    2004
<210> 80
<211> 3555
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 3295764CB1
<400> 80
gaccgtgggc gagtccagaa cgtcctggcc ttacagggag aaggcgtcac tcgcggttac 60
aagtgcctga ccctcactcc agttggcgga ggaggagaag gaaggggccg ggccgggtcc 120
ceteceteg egeceeggat ggatqtqcc qqceqqtqt eteqqeqqq ggeggegg 180
geggcactg tgctcctgcg gaccgctcgg gtccgtcgcg aatgctggtt cttgccgacc 240 gcgctgctct gcgcctacgg cttcttcgcc agcctcaggc cgtccgagcc cttcctgacc 300 ccgtacctgc tggggccgga caagaacctg accgagaggg aggtcttcaa tgaaatttat 360
ccagtatgga cttactctta cctggtgcta ctgtttcctg tgttccttgc cacagactac 420
ctccgttata aacctgttgt tctactgcag gggctcagcc ttattgttac atggtttatg 480 ctgctctatg cccagggact gctggccatt caatttctag aatttttta tggcatcgcc 540
acagccactg aaattgccta ttactcttat atctacagtg tggtggacct gggcatgtac 600
cagaaagtca caagttactg tcgaagtgcc actttggtgg gctttacagt gggctctgtc 660 ctagggcaaa tccttgtctc agtggcaggc tggtcgctgt tcagcctgaa tgtcatctct 720
cttacctgtg tttcagtggc ttttgctgtg gcctggtttt tacctatgcc acagaagagc 780
ctcttctttc accacattcc ttctacctgc cagagagtga atggcatcaa ggtacaaaat 840
ggtggcattg ttactgacac cccagcttct aaccaccttc ctggctggga ggacattgag 900
tcaaaaatcc ctctaaatat ggaggagcct cccgtggagg aaccggaacc caagccagac 960
egteteettg tattgaaagt actatggaat gattteetga tgtgetaete etetegeeet 1020
cttctctgct ggtctgtgtg gtgggccctc tctacctgtg gctattttca agttgtgaac 1080 tacacacagg gcctgtggga gaaagtgatg ccttctcgct atgctgctat ctataatggt 1140
ggcgtggagg ccgtttcaac cttactgggt gctgttgctg tgtttgcagt tggttatata 1200
aaaatateet ggteaacttg gggagaaatg acattatete tetttetet eetgattget 1260
gctgcagtgt atatcatgga cactgtgggt aacatttggg tgtgctatgc atcctatgtt 1320
gtottcagaa toatctacat gttactcatc acqataqcaa cttttcaaat tgctgcaaac 1380
ctcagcatgg aacgctatgc cctagtattt ggtgtaaata ccttcattgc cctggcactg 1440
cagacgctgc tcactctaat tgtggtagat gccagtggcc ttggattaga aattaccact 1500
cagtttttga tctatgccag ttattttgca ctcatcgctg tggttttcct ggccagtggt 1560
```

```
gcagtcagtg ttatgaagaa atgtagaaag ctggaagatc cacaatcaag ttctcaagta 1620 accacttcat aatatactgc tgaagggctt cttcttatag caagaactct gcacagcaac 1680 tgcctggatg tatttgattt tttaaagcgt agacatatat ttatgaatgt gcattcttg 1740
acttcacage agecaettga etaatacett gtgttceggg aataacatga tactattcag 1800
aggagccaga agtaaagttt atttcatgga ttatttatga gagctaattt aaggatgact 1860
ttttttctga ttcaaaagtg aacttgattt taaaaaccag tcaagagcaa tcaaagcagc 1920 acatggtgtt gtatacttca ttagcaagtg agtttggtgt tttataggtc acatatgtct 1980
gtatctactt agccagatgc ttggcctggt gggaccaggg ctccacagag gccacaaaat 2040 gttgcaagtc atgatggatg gaaatatgtt ctaacagcat ctgcctctat tcaatttaat 2100
tettatttet gtgttaetea tgtacattgg tetttetaea tagttattet ateaetggea 2160
atatttgttc tggtttagtg ttctgtattt taaggtgtac gtatcatttc taattttaag 2220
ttattttaaa aaaattcatc atatgaatgt tcttggttcc cattgtgacg attatttatt 2280 tctgtaaaat ttgtttagaa gtacgttttt gcattattca tatgcttccc agagaagctc 2340
atttagttag aaaataaggc aagttttgaa gcctgctaaa tgaagagact taagaaagct 2400
taaggtacgc ttgcttgtct ttaaatcttc aatatgaagg actattaatt ccaagattaa 2460 aagttcatat ataggctaaa gatgtaacta ggccatttgt atttgtattc ccttttattt 2520
ccaaaataaa atgaaaaatc ttttttaat aatttcatcc ctatttatag tttttatatt 2580
aatttgttt tcttatccaa gtaaagatgt caataggaat tgcattagtc caaggccttt 2640 ttcataaact gagcctcttt tcaattattt caatgggaca ggaactagga tagatgtgat 2700
tectgeattt ttttacetta aatetgeett tgtttetaaa ggtagateat ettgaatatt 2760
tgcttaaaat tgctagtgat ttcattacca agttacttga aaaaatgttc tatatgcatt 2820
taattetgaa ateagtetae caaggggetg etagtatatg teagacatga aaactatttt 2880 aaagetgaet ttgttgeett atettgaaaa gaatetagat aggtgettt aaetggggta 2940
ttaacttttt tagaatgaca cagctgaaca gtgttaataa tagtgtgtca agattgcaaa 3000
gtcgacatac tcatttggtt taagcaggaa tcctagaagc aaatggatgg ggataagaat 3060 aggtcatttt ctattcacca tcctttacta ttaagggaaa ggaaaagaac actagctaag 3120
gaagggaaag ggaagtgata tcataaaagt agcaaccttc attttacatt ctgtctgttt 3180 ttcttttttt gctttgttt gtttgtgcta atttgggaat tgtgtactcc gaaacaagta 3240 gaaaagtgct gtttgaggga ttttattaaa tcttttttta atggaatgtg gtacaaatg 3300
ttcatgttac caaagcaata tttccctgga atttaattca aagtttgtgg catacaacct 3360
gagcettte ttatataaga caagaatatg tteacatett ggtatgtgge catatttata 3420 gaatgetgaa eteaatgtge aagttgtaet gtatgeagtt ttgtaaataa gtgaaaataa 3480
ttttacagta tattt
                                                                                             3555
<210> 81
<211> 1293
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 3438320CB1
<400> 81
atatgececa egegggaete atactaegtt teeegtgaac aegtgeagte caaaceeege 60
ccctgatatt tatctcagtg gacggtggcc ggaaaaggac aatggtttcc atgtcagcgg 120 ataaacgctc tcccctcggc tcccggacgc gacggaggtc gtagtagtag tgagtacgtg 180
ctgaggagca aaggagtaac caagagatcc agtgaccgac agagcaagag ccatgccgcg 240
ccggggcctg gtggctgggc cagacttgga gtattttcag cgtcgctatt tcacgccggc 300
ggaggtggcc caacataaca ggcccgaaga cctctgggta tcttacctgg gacgcgtgta 360
cgacctaacg tcattggcac aggaatacaa ggggaacctq ctqctqaaac ccatcgtgga 420
agttgcaggc caggatatca gccactggtt tgatccaaag accagagaca tccgcaagca 480
catagatccg ctgaccggct gcctgaggta ctgcaccccg cggggccgct ttgtgcacgt 540 tccgcctcag ctgccctgtt cggactggc caacgatttt gggaagccct ggtggcaggg 600
gtcgtattat gaggtggggc ggctgtctgc caagacccgg agcatccgca tcattaacac 660
gctcacgtcg caggagcaca cactggaggt gggggttctg gagtccatat gggaaatcct 720 acaccgctat ctcccctata actcacatgc tgccagctac acgtggaaat atgaagggaa 780
gaacctgaac atggatttta ccctggaaga gaatgggatc cgggatgagg aggaagaatt 840
tgactatctc agtatggacg gtacacttca cacacctgca atacttctgt acttcaatga 900 tgatctcacg gagttgtagg caaggagatg tacactcgtg tagactcaag acgtatttcg 960
agtttggctt tttctgtgcc ttgaggaaaa gtggtggggc cgaggggtgc ctggacccag 1020
atctccactc ctctccagga gctagcctgt gcccttctga agtgtaaagg ccctattccc 1080
tgccttcatt acagtttgct ctgagaaaat tagtgaatta atctttggga atgatacaag 1140 aagatcaagt accttggttt agggagatgt agaagaggat agtcagagtt caggcagaac 1200
tgtttgatag ttaagagaga gtagttctac aggggtgagg gatggaagga cttttttggc 1260
aatgatggaa atgagatgtc tgcaggagat ggg
                                                                                             1293
```

```
<210> 82
<211> 1489
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 3986488CB1
cggggctggc gggcggcgct cttctacggg acctgctcct tcctcatcgt gcttgtcaac 60
aaggegetge tgaccaecta eggttteeeg teaccaattt teettggaat tggacagatg 120
gcagccacca taatgatact atatgtgtcc aagctaaaca aaatcattca cttccctgat 180
tttgataaga aaatteetgt aaagetgttt eetetgeete teetetaegt tggaaaceae 240
ataagtggat tatcaagcac aagtaaatta agcctaccga tgttcaccgt gctcaggaaa 300
ttcaccattc cacttacctt acttctggaa accatcatac ttgggaagca gtattcactc 360 aacatcatcc tcagtgtctt tgccattatt ctcggggctt tcatagcagc tgggtctgac 420
cttgctttta acttagaagg ctatattttt gtattcctga atgatatctt cacagcagca 480
aatggagttt ataccaaaca gaaaatggac ccaaaggagc tagggaaata cggagtactt 540
ttctacaatg cctgcttcat gattatccca actcttatta ttagtgtctc cactggagac 600
ctgcaacagg ctactgaatt caaccaatgg aagaatgttg tgtttatcct acagtttctt 660
ctttcctgtt ttttggggtt tctgctgatg tactccacgg ttctgtgcag ctattacaat 720
tcagccctga cgacagcagt ggttggagcc atcaagaatg tatccgttgc ctacattggg 780 atattaatcg gtggagacta cattttctct ttgttaaact ttgtagggtt aaatatttgc 840
atggcagggg gcttgagata ttccttttta acactgagca gccagttaaa acctaaacct 900
gtgggtgaag aaaacatctg tttggatttg aagagctaaa gagtctgcag caggattgga 960 gactgacttg tgactgcggg ctggggggc attcccagta ggaatgtgaa gccagaggtt 1020
teggattegt gacatecace ecetgggeaa gtgagageat etgeaaaatg caaagagaae 1080 taceteatat geaggatgag ceaatggeag teteaagaaa tgtaeteggg egacacetta 1140 cetgtggaaa geaaatettt teaaaataag eeactgggae teggtaggtg gageeecage 1200
tgctcttcta gggacctatg gggccttcgt ggcatctctg tgctgtgtgc tggggaggag 1260 gttgatgtaa tggtgactct tttctgatca gcaccttggc cgtgattccc aaggtcccag 1320 ccaaagcaaa gggccagttg tttcagttta aacagacatg tctttagtct aataaaatta 1380
gttaactgcc agtaaagtta tttgttagct ttgatgaaag ctatgttggt atctttccct 1440
aatcatcaaa gtaaataaaa aatcatttct atgtaaaaaa aaaaaaaaa
<210> 83
<211> 927
<212> DNA
<213> Homo sapiens
<221> misc_feature
<223> Incyte ID No: 4378816CB1
ctccctgctt tcctctgccg catggtcctg ggccgttggc gtcggaagcc tgaagcatgg 60
gegetgagtg ggagetgggg geegaggetg geggtteget getgetgtge geegegetge 120 tggeggggg etgegeetg ggeetgege tgggeegegg geggaeegeg 180 gggegeteat etggetetge taegaegege tggtgeaett eggetggaa ggeeettttg 240
tetaettete titagtagga aacetteeaa atteegatee etteatteet tettatega 300
aagaatatgg caaagctgat gcaagatggg tttattttga tccaaccatt gtgtctgtgg 360 aaattctgac cgtcgccctg gatgggtctc tggcattgtt cctcatttat gccatagtca 420 aagaaaaata ttaccggcat ttcctgcaga tcaccctgtg cgtgtgcgag ctgtatggct 480
getggatgae ettectecca gagtggetea eeagaageee caaceteaae accageaaet 540
ggctgtactg ttggctttac ctgtttttt ttaacggtgt gtgggttctg atcccaggac 600
tgctactgtg gcagtcatgg ctagaactca agaaaatgca tcagaaagaa accagttcag 660
tgaagaagtt tcagtgaact ttcaaaacca taaacaccat tatctaactt catgaaccag 720 aatgaatcaa atcttttgt ttggccaaaa tgtaatacat tccagtctac actttgttt 780
tgtattgttg ctcctgaaca acctgtttca aattggtttt aaggcgacca gttttcgttg 840
tattgttgtt caattaaatg gtgatatagg gaaaagagaa caaatttgaa tttgtaataa 900
taaaatgttt aattatacaa aaaaaaa
                                                                                            927
<210> 84
<211> 970
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 4797137CB1
<400> 84
ggatgcagca gagaggagca gctggaagcc gtggctgcgc tctcttccct ctgctgggcg 60 tcctgttctt ccagggtgtt tatatcgtct tttccttgga gattcgtgca gatgcccatg 120
tccgaggtta tgttggagaa aagatcaagt tgaaatgcac tttcaagtca acttcagatg 180
tcactgacaa gcttactata gactggacat atcgccctcc cagcagcagc cacacagtat 240 caatatttca ttatcagtct ttccagtacc caaccacagc aggcacattt cgggatcgga 300
tttcctgggt tggaaatgta tacaaagggg atgcatctat aagtataagc aaccctacca 360
taaaggacaa tgggacattc agctgtgctg tgaagaatcc cccagatgtg caccataata 420 ttcccatgac agagctaaca gtcacagaaa ggggttttgg caccatgctt tcctctgtgg 480
cecttette cateettgte tttgtgeeet cageegtggt ggttgetetg etgetggtga 540
gaatggggag gaaggctgct gggctgaaga agaggagcag gtctggctat aagaagtcat 600
ctattgaggt ttccgatgac actgatcagg aggaggaaga ggcgtgtatg gcgaggcttt 660 gtgtccgttg cgctgagtgc ctggattcag actatgaaga gacatattga tgaaagtctg 720
tatgacacaa gaagagtcac ctaaagacag gaaacatccc attccactgg cagctaaagc 780 ctgtcagaga aagtggagct ggcctggacc atagcgatgg acaatcctgg agatcatcag 840 taaagacttt aggaaccact tattattga ataaatgttc ttgttgtatt tataaactgt 900
tcaggaagtc tcataagaga ctcatgactt cccctttcaa tgaattatgc tgtaattgaa 960
tgaagaaatc
                                                                                           970
<210> 85
<211> 594
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5470806CB1
gacaggatgg cttcccttcg cctgttcctc ctctgcctcg ctgtactgat atttgcgtct 60
gaagctggcc ctgggggtgc tggagaatcc aagtgtcctc tgatggtcaa agtcctggat 120 gctgtccgag gcagccctgc tgttgatgtg gccgtgaaag tgttcaaaaa gactgcagac 180
ggaagctggg agccgtttgc ctctgggaag accgccgagt ctggagagct gcacgggctc 240
accacagatg agaagttcac ggaaggggtg tacagggtag aactggacac caaatcgtac 300 tggaaggctc ttggcatttc cccattccat gaatacgcag aggtggtttt cacagccaat 360
gactetggte ategecaeta caccategea gecetgetea gecegtaete etacageaec 420
actgctgtcg tcagtaaccc ccagaactga gggacccagc ccagtaggac caggatcttg 480 ccaaagcagt agcttcccat ttgtactgaa acagtgttct tgctctataa accgtgttag 540
caactoggga agatgcogtg aaacgatott attaaaccac otgtgatgco aaaa
                                                                                            594
<210> 86
<211> 618
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 5473242CB1
gtgttgactc gcaacctcag gaacagacac catggtgcac ctaactgatg ctgagaaggc 60
tactgttaat ggcctgtggg gaaaggtgaa ccctgttgaa attggcgctg agtcccttgc 120
cagtetgetg attgtetace ettggaccca gaggtacttt tetaaatttg gggacetgte 180
ctctgtctct gctatcatgg gtaaccccca ggtgaaggcc catggcgaaa aggtgataaa 240 cgccttcgat gatggcctga aacacttgga caacctcaag ggcacctttg ccagcctcag 300
tgaactccac tgtgacaagc tgcatgtgga tcctgagaac ttcaggctcc tgggcaatat 360
gattgtgatt atgatgggcc accacctggg caaggaattc accccgagtg cacaggctgc 420 cttccagaag gtggtggctg gagtggccag tgccctggct cacaagtacc actaaacctc 480
ttttcctgct cttgtctttg tgcaatggtc aattgttccc aagagagctt ctgtcagttg 540
ttgtcaaaat gacaaagacc tttgaaaatc tgtcctacta attaaagcat ttggttcaag 600
tgttctgttg agataccc
                                                                                            618
```